

## The neural time course of semantic ambiguity resolution in speech comprehension

Article (Accepted Version)

MacGregor, Lucy J, Rodd, Jennifer M, Gilbert, Rebecca A, Hauk, Olaf, Sohoglu, Ediz and Davis, Matthew H (2019) The neural time course of semantic ambiguity resolution in speech comprehension. *Journal of Cognitive Neuroscience*. pp. 1-23. ISSN 0898-929X

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/88042/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

### **Copyright and reuse:**

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

The neural time course of semantic ambiguity resolution in speech comprehension

Lucy J. MacGregor<sup>1\*</sup>

Jennifer M. Rodd<sup>2</sup>

Rebecca A. Gilbert<sup>1</sup>

Olaf Hauk<sup>1</sup>

Ediz Sohoglu<sup>1</sup>

Matthew H. Davis<sup>1</sup>

<sup>1\*</sup> MRC Cognition and Brain Sciences Unit, University of Cambridge

15 Chaucer Road, Cambridge, CB2 7EF

Email: [lucy.macgregor@mrc-cbu.cam.ac.uk](mailto:lucy.macgregor@mrc-cbu.cam.ac.uk)

Tel: +44 (0)1223 273708

Fax: +44 (0)1223 359062

<sup>2</sup>Department of Experimental Psychology, University College London

Semantically ambiguous words challenge speech comprehension, particularly when listeners must select a less frequent (subordinate) meaning at disambiguation. Using combined MEG and EEG, we measured neural responses associated with distinct cognitive operations during semantic ambiguity resolution in spoken sentences: (i) initial activation and selection of meanings in response to an ambiguous word, and (ii) sentence reinterpretation in response to subsequent disambiguation to a subordinate meaning. Ambiguous words elicited an increased neural response approximately 400 to 800 ms after their acoustic offset compared to unambiguous control words in left fronto-temporal MEG sensors, corresponding to sources in bilateral fronto-temporal brain regions. This response may reflect increased demands on processes by which multiple alternative meanings are activated and maintained until later selection. Subsequent, disambiguating words heard after an ambiguous word were associated with marginally-increased neural activity over bilateral temporal MEG sensors, and a central cluster of EEG electrodes, which localised to similar bilateral frontal and left temporal regions. This later neural response may reflect effortful semantic integration, or elicitation of prediction errors that guide reinterpretation of previously-selected word meanings. Across participants, the amplitude of the ambiguity response showed a marginal positive correlation with comprehension scores, suggesting that sentence comprehension benefits from additional processing around the time of an ambiguous word. Better comprehenders may have increased availability of subordinate meanings, perhaps due to higher-quality lexical representations and reflected in a positive correlation between vocabulary size and comprehension success.

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

## Introduction

Most common words are semantically ambiguous (for a review, see Rodd, Gaskell, & Marslen-Wilson, 2002), such that their meaning depends on context. For example, “ace” can refer to a playing card or a tennis serve that an opponent is unable to return. Thus, the ability to make sense of – resolve – ambiguity is a fundamental part of speech comprehension. When listeners (or readers) encounter an ambiguous word (e.g., “ace”) semantic priming studies suggest that they automatically activate the multiple meanings of that word in parallel (irrespective of context) but within a few hundred milliseconds, settle on a single preferred meaning (Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Swinney, 1979). Initial meaning selection operates on the information available at that time (Cai et al., 2017; Duffy, Morris, & Rayner, 1988; Rodd, Cutrin, Kirsch, Millar, & Davis, 2013; for a review, see Vitello & Rodd, 2015) which will be particularly challenging if disambiguating context is absent or delayed until after the ambiguous word. If subsequent context supports a subordinate (less frequent, thus more unexpected) meaning, then a later process of reinterpretation is often necessary for accurate comprehension.

Individual differences in comprehension success have been associated with abilities at accessing, selecting and reinterpreting ambiguous word meanings (Gernsbacher, Varner, & Faust, 1990; Henderson, Snowling, & Clarke, 2013; Szabo Wankoff & Cairns, 2009). Damage to the anterior temporal lobe, a region known to be associated with semantic processing in general (Patterson, Nestor, & Rogers, 2007), has been shown to impair the processing of ambiguous word meanings (Zaidel, Zaidel, Oxbury, & Oxbury, 1995), but it is still unclear how variation in comprehension ability relates to variation in the associated neural processes. The aim of the current study is to understand the neural mechanisms that support two stages of successful ambiguity resolution: initial meaning activation/selection and subsequent reinterpretation, and to explore the relationship between behavioural and neural responses to ambiguity.

The cortical network supporting ambiguity resolution in sentences was first reported in a functional MRI study by Rodd and colleagues (Rodd, Davis, & Johnsrude, 2005). Listeners were

presented with high ambiguity sentences containing multiple ambiguities (e.g. “there were DATES and PEARS on the kitchen table”) and the associated BOLD activation was contrasted with that produced by low ambiguity control sentences (e.g. “there was beer and cider on the kitchen shelf”). Additional activation during comprehension of high ambiguity sentences was observed in bilateral inferior frontal gyrus (IFG), particularly in pars triangularis and opercularis, and in left posterior temporal regions, including posterior middle temporal gyrus (pMTG), inferior temporal gyrus (pITG) and fusiform. These activations were observed in the absence of explicit awareness of the ambiguities and when listeners were given no explicit task, suggesting involvement of these regions when comprehension occurs automatically as in natural speech comprehension. This basic observation that semantic ambiguity resolution involves fronto-temporal regions is now well established, having been replicated using functional MRI for spoken (Rodd, Johnsrude, & Davis, 2012; Rodd, Longe, Randall, & Tyler, 2010; Tahmasebi et al., 2012; Vitello, Warren, Devlin, & Rodd, 2014) and written (Mason & Just, 2007; Zempleni, Renken, Hoeks, Hoogduin, & Stowe, 2007) sentences and shown to have a consistent localisation across individuals (Vitello et al., 2014). This fronto-temporal response to ambiguity has proved useful in translational work, for example as a neural marker of residual semantic processing of speech at different levels of sedation (Davis et al., 2007) and as evidence for intact speech comprehension which has prognostic value for patients diagnosed as being in a vegetative state (Coleman et al., 2009; Coleman et al., 2007).

However, attempts to attribute specific cognitive operations like initial meaning activation/selection and subsequent reinterpretation to distinct cortical regions have been less successful. One experimental approach has been to compare neural responses to sentences containing ambiguous words with varying meaning frequencies; such sentences are expected to load on different processes in ambiguity resolution. For example, initial meaning selection is assumed to be more difficult for sentences containing ambiguous words with meanings that have similar frequencies (balanced) than for words with a more dominant meaning (biased). Conversely, reinterpretation is assumed to be more difficult or more likely when sentences are disambiguated to

a subordinate (less frequent and therefore less expected) meaning. In this way, BOLD responses due to differences in meaning frequency can be related to processes at the time of ambiguity (initial meaning activation/selection) or disambiguation (subsequent reinterpretation). Using this approach, responses to subordinate meanings have been attributed to reinterpretation processes in the left (Vitello et al., 2014) or bilateral (Mason & Just, 2007; Zempleni et al., 2007) IFG, sometimes extending into superior and middle frontal areas (Mason & Just, 2007). However, posterior MTG/ITG has also been implicated in reinterpretation, with studies observing greater activation for subordinate meanings in left (Vitello et al., 2014) or bilateral (Zempleni et al., 2007) posterior temporal regions, though null results are also reported (Mason & Just, 2007). Initial meaning selection has also been associated with responses in left IFG (Mason & Just, 2007) but other studies have failed to observe greater activation for balanced compared to biased ambiguous words and hence evidence for selection processes is currently lacking (Vitello et al., 2014).

An alternative approach to separating neural responses during initial meaning selection from those involved in subsequent reinterpretation has explored differences used the timing of fronto-temporal responses. Rodd and colleagues (Rodd et al., 2012) used a rapid fMRI acquisition sequence to measure the time course of the BOLD response to ambiguous sentences in which the timing of disambiguation was varied. They assumed that additional BOLD responses associated with reinterpretation (relative to unambiguous control sentences) would occur later for ambiguous sentences in which disambiguation occurred after an additional delay. Hence, they contrasted delayed disambiguation sentences, like “The ecologist thought that the PLANT by the river should be closed down” with immediate disambiguation sentences, like “The scientist thought that the FILM on the water was from the pollution” (AMBIGUOUS and disambiguation words highlighted). BOLD responses to immediate and delayed ambiguity resolution showed differences in timing in Left IFG and in posterior temporal areas (fusiform, pITG and pMTG) consistent with reinterpretation. Furthermore, BOLD responses were also observed in the IFG for sentences in which the disambiguating information occurred prior to the ambiguous word (“The hunter thought

1 that the HARE in the field was actually a rabbit”). Since these sentences should not require  
2  
3  
4 reinterpretation Rodd and colleagues concluded that the IFG is also involved in meaning selection.

5  
6 Taken together, an emerging picture of the differential contribution of inferior frontal and  
7  
8 posterior temporal brain regions to semantic ambiguity resolution is that meaning selection may be  
9  
10 underpinned by IFG and reinterpretation by IFG and posterior temporal areas together. However,  
11  
12 there is a lack of consistent findings in relevant experiments perhaps due to the challenge of  
13  
14 associating a slow BOLD response, which has a rise-time of around 5 seconds (Boynton, Engel,  
15  
16 Glover, & Heeger, 1996; Josephs & Henson, 1999) with distinct neurocognitive processes that  
17  
18 operate over a shorter time period. This leads to two problems. First, during the comprehension of a  
19  
20 single sentence lasting less than 5 seconds, the measured BOLD response to different  
21  
22 neurocognitive events will inevitably overlap making it difficult to tease apart initial meaning  
23  
24 activation/selection and subsequent reinterpretation. Second, given that meaning selection is  
25  
26 thought to occur within a few hundred milliseconds (Seidenberg et al., 1982; Swinney, 1979), the  
27  
28 BOLD response may be insensitive to a transient neural response.  
29  
30  
31  
32  
33

34 Several studies have utilised more temporally-sensitive measures of cognition to investigate  
35  
36 the processing of ambiguous words. During natural reading, the duration of fixation times have  
37  
38 been shown to be longer for ambiguous words in the absence of biasing context compared to  
39  
40 unambiguous controls (Frazier & Rayner, 1990; although for evidence that reading times for  
41  
42 ambiguous words with biased meanings do not differ from unambiguous controls, see Duffy et al.,  
43  
44 1988; Rayner & Duffy, 1986). ERP studies with word-by-word presentation have shown a sustained  
45  
46 frontal negativity for ambiguous words presented in a semantically neutral context compared to  
47  
48 unambiguous words (Hagoort & Brown, 1994), and for ambiguous words in a semantically neutral,  
49  
50 but syntactically constraining context compared to unambiguous controls (Federmeier, Segal,  
51  
52 Lombrozo, & Kutas, 2000; C. L. Lee & Federmeier, 2006, 2009, 2012). These findings suggest that  
53  
54 processing of ambiguous words is more effortful than processing words with single meanings. ERP  
55  
56 studies using word-by-word visual presentation have also looked for effects potentially associated  
57  
58  
59  
60

1 with reinterpretation (Gunter, Wagner, & Friederici, 2003; Hagoort & Brown, 1994). In these  
2  
3 studies, N400 responses have been observed in response to disambiguating words that resolve an  
4  
5 ambiguity to its subordinate meaning. However, these studies did not control for both the  
6  
7 presence/absence of ambiguity and the word form itself. Hence, differences in word form and  
8  
9 meaning might also be responsible for these neural effects.  
10  
11

12  
13 In the present study we used combined magnetoencephalography (MEG) and  
14  
15 electroencephalography (EEG) which provides the temporal resolution required to distinguish  
16  
17 neural responses at different time points during sentences and to relate these responses to distinct  
18  
19 neurocognitive processes. Our volunteers listened to spoken sentences (see Figure 1A) that  
20  
21 manipulated the presence/absence of an ambiguous word (AMBIGUITY) and subsequent  
22  
23 disambiguation (*disambiguation*, e.g., “The man thought that one more ACE/SPRINT might be  
24  
25 enough to win the *tennis/game*.”). These sets of sentences enable us to specify conditions and time-  
26  
27 points in which we expect either initial meaning access and selection or reinterpretation to occur.  
28  
29  
30

31  
32 In the absence of biasing context an ambiguous word (ACE) should require additional  
33  
34 meaning access and selection processes relative to a matched, unambiguous control word  
35  
36 (SPRINT). This comparison of sentences with and without an ambiguous word (i.e. the main effect  
37  
38 of ambiguity) provides the first experimental contrast in our study. Neural activity during and after  
39  
40 the ambiguous word will reflect processes involved in initial meaning activation and selection that  
41  
42 are more strongly taxed by ambiguous than control (unambiguous) words. These processes should  
43  
44 occur prior to subsequent context words that drive reinterpretation.  
45  
46  
47

48  
49 Given that the words that precede the ambiguous word are relatively uninformative, initial  
50  
51 meaning access and selection should result in most listeners settling on the dominant (playing card)  
52  
53 meaning of the ambiguous word. The subsequent presentation of a sentence-final word (*tennis*) that  
54  
55 is incompatible with the dominant meaning of ACE disambiguates the ambiguous word to its  
56  
57 subordinate meaning. For listeners to avoid misinterpretation, a resource-demanding  
58  
59 reinterpretation processes should therefore be triggered by the sentence final word (*tennis*) but not  
60



by a final word (*game*) that is consistent with both meanings (Duffy et al., 1988; Kambe, Rayner, & Duffy, 2001; Rodd, Johnsrude, & Davis, 2010). Since this reinterpretation process will only occur if the sentence-final word (*tennis*) occurs in a sentence that contains the ambiguous word (ACE), the neural correlates of reinterpretation can be detected using the interaction between ambiguous words and subordinate reinterpretation, time-locked to the sentence-final word.

For both these meaning access/selection (main effect) and reinterpretation (interaction) contrasts, we measured evoked MEG/EEG responses relative to the offset of the critical words. This is a time point at which listeners have heard sufficient phonetic information to recognise the words and are therefore engaged in meaning processes. We used an active comprehension task on non-critical trials during MEG/EEG scanning to ensure attentive listening throughout without contaminating neural measures obtained during critical trials.

In addition to our analyses of main effects and interactions, we were also interested in relating neural responses to individual differences in sentence comprehension. We therefore administered a post-scanning behavioural task to provide a trial-by-trial measure of the comprehension of critical sentences that required reinterpretation of an ambiguous word. We were interested in whether more successful ambiguity resolution would be associated with greater neural engagement or reduced processing effort. We were also interested in whether there was a relationship between comprehension and verbal and non-verbal abilities (as measured using standard vocabulary and fluid reasoning tests).

Figure 1 here

## Materials and Methods

### *Stimuli*

Sets of 80 spoken sentences were constructed according to a two-by-two factorial design in which we manipulated (1) the presence/absence of an ambiguous word (Ambiguity: ambiguous vs.

control) and (2) the presence of one of two sentence-final words, which in the ambiguous sentences either disambiguated the ambiguous word so it resolved to a subordinate meaning, or left it unresolved (Disambiguation: resolved vs. unresolved). Since identical sentence-final words also completed the unambiguous control sentences we also use the terms resolved/unresolved to refer to the equivalent control conditions (see Figure 1A, Table 1. Ambiguous words occurred mid-sentence after a neutral context that did not bias interpretation towards either meaning of the ambiguous word (mean word offset of 1423 ms after sentence onset; see Figure 1B), and were followed by additional neutral context words. In the ‘ambiguous-resolved’ sentences the sentence-final word disambiguated the ambiguous word towards a subordinate meaning (mean word onset and offset were 1068 ms and 1506 ms after the offset of the ambiguous word; see Figure 1B). In the ‘ambiguous-unresolved’ sentences, the sentence-final word were necessarily more general so that both meanings of the ambiguous word remained plausible. Identical sentence-final words also completed the control unambiguous sentences. Sentence transcriptions and stimulus properties can be downloaded from: <https://osf.io/3jhtb/>.

Table 1 here

The critical 80 ambiguous and 80 unambiguous control words were matched on mean frequency of occurrence, number of syllables, and number of phonemes (Baayen, Piepenbrock, & Gulikers, 1995). Sentence-final words, that in the ambiguous sentences either did or did not resolve the ambiguities were also matched on the same factors (Table 2).

Table 2 here

Analysis of a large database of meaning dominance ratings for single ambiguous words (Gilbert, Betts, Jose, & Rodd, 2017) created using standard word association methods (Twilley, Dixon, Taylor, & Clark, 1994) confirmed that the ‘ambiguous resolved’ condition sentences utilised the subordinate meaning of the ambiguous words, with the exception of a small number of sentences (mean dominance = 0.23, SD = 0.21, max = 0.76, min = 0). The ‘ambiguous resolved’ condition sentences were also tested using the word association method to ensure that

disambiguation to the subordinate meaning occurred only at the sentence-final word (not earlier):

Participants who did not take part in the MEG experiment were presented with the ‘ambiguous resolved’ condition sentences without the final word, followed by the isolated ambiguous word, and asked to generate a word that was related to the ambiguous word as used in the sentence.

Dominance ratings of the ambiguous words in context were comparable to those taken from the database of isolated ambiguous words (mean dominance = 0.25, SD = 0.16, max = 0.53, min = 0).

Meaning dominance ratings can be downloaded from: <https://osf.io/3jhtb/>.

The 80 ambiguous words and their matched unambiguous control words were used to create 80 stimulus sets. Within each set there were two lead-in contexts (e.g., “The man knew...” and “The woman hoped...”), which were crossed with the ambiguous/control words and the sentence-final ambiguity-resolving/unresolving words, thus resulting in eight stimulus versions. For each set, the eight versions were separated into two lists – list A and list B, each containing one sentence from each of the four conditions such that each ambiguous/control word and sentence-final word occurred twice, but following a different lead-in context. Participants heard stimuli from either list A or list B (320 stimuli in total) which meant that although they heard each ambiguous word twice – in a resolved and an unresolved sentence, each followed a different lead-in context (see Figure 1A and Table 1 for examples of stimulus sets heard by one participant).

The stimuli were spoken by a native speaker of Southern British English (author MHD) and digitally recorded (44.1 KHz sampling rate) in a sound-proofed booth. For each stimulus set, all eight versions of the sentences were recorded, then six segments were extracted from the recordings, corresponding to the lead in portion (two versions), the target word (ambiguous, unambiguous) plus surrounding words, and the sentence-final word plus surrounding words (see shading in Figure 1A). The six segments were then concatenated to make the eight sentence versions, which were carefully checked to ensure no splices were audible. The procedure of splicing and then recombining segments meant that across conditions, the critical sections of each sentence (e.g. ambiguous word, disambiguation) were acoustically identical. The exact point for splicing was

1  
2 chosen to ensure that the recombined stimuli sounded natural (e.g., by selecting silent periods  
3 during plosives). Stimuli were normalised within and between-conditions for RMS amplitude using  
4 Praat software (from <http://www.praat.org>).  
5  
6  
7

8  
9 In addition to the experimental stimuli, 20 sets of filler sentences were constructed with  
10 similar lexico-syntactic structures and properties as the experimental stimuli. There were four  
11 sentence versions per set in which the ambiguous/control words were crossed with the sentence-  
12 final ambiguity-resolving/unresolving words (80 fillers in total); as with the experimental stimuli  
13 the ambiguous/control words and sentence-final words occurred twice but with a different lead in  
14 for each repetition. RMS amplitudes of the fillers were adjusted to match the mean RMS amplitude  
15 of the experimental files. Participants heard all filler sentences. For each of the filler sentences,  
16 probe words were selected for visual presentation in the relatedness judgement task, which was  
17 included to probe for comprehension and to ensure attentive listening. Probe words were either  
18 strongly related (50% of probes) or unrelated (50% of probes) to the meaning of the sentence  
19 meaning. The probes were never related to the unintended meaning of the ambiguous words.  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35

36 *Cloze probability test*  
37  
38

39 Following a suggestion from a reviewer we ran a sentence completion test on our four experimental  
40 sentence types to test whether there were differences in cloze probability across the four conditions.  
41 Data were collected from 77 participants (aged 20-39 years, born and residing in the UK, who had  
42 learned English as their first language and had no hearing difficulties) over the internet using  
43 jsPsych (de Leeuw, 2015) and JATOS (Lange, Kuhn, & Filevich, 2015), following recruitment via  
44 Prolific (Palan & Schitter, 2017; Peer, Samat, Brandimarte, & Acquisti, 2015). Data from five  
45 participants were excluded (see below) and were replaced in order to meet our a priori goal of  
46 analysing 72 data sets (giving us a cloze probability resolution for each item of 1.4%).  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56

57 The same sets of 80 sentences from the MEG study were used in this test, except that the  
58 final words of each sentence were not presented. Thus for each of the 80 experimental items, there  
59  
60

1  
2 were four possible sentences created by crossing the two lead-in versions with the two key words  
3  
4 (ambiguous or control). To avoid excess stimulus repetition each participant was tested on only two  
5  
6 out of four sentence variants (i.e. they heard each lead-in version only once, with one variant  
7  
8 presented with an ambiguous word and the other with a control word; 160 experimental item trials  
9  
10 in total). We counterbalanced whether ambiguous or control words were presented first for specific  
11  
12 items, and which lead-in variant was paired with an ambiguous word resulting in four experimental  
13  
14 versions. While we aimed at testing 18 participants in each of the four experimental versions, due to  
15  
16 accidental over-recruitment in one version we collected data from 19 participants in one version, 18  
17  
18 in two versions and 17 in another version.  
19  
20  
21

22  
23 Participants were told that they would hear sentences in which the ending has been cut off,  
24  
25 and their task was to complete the sentence with the word or words that first come to mind. In each  
26  
27 trial, a spoken sentence was presented up until the splice point at which the resolved and unresolved  
28  
29 sentences diverged acoustically (see Figure 1B, i.e. a silent period between the key word and the  
30  
31 sentence-final word). This allowed us to avoid presenting co-articulatory or other cues that could  
32  
33 constrain or bias listeners' choice of sentence-final words. However, because the splice point often  
34  
35 occurred two or more words before the end of the sentence we also presented the remaining words  
36  
37 before the sentence-final word as written text. For example, for the item 'ace', listeners would hear :  
38  
39 "The man knew that one more ACE might be enough" (lead-in 1, ambiguous/control key word) and  
40  
41 see: "to win the...", followed by a text entry box for a sentence completion response. For splice  
42  
43 points occurring in the middle of a word these words were also presented at the start of the text  
44  
45 segment to avoid confusion. Splice points occurred at the same place for all four sentences for each  
46  
47 item and hence the text presented on the screen the same for all four versions of each sentence. In  
48  
49 addition to the cloze task, as in the MEG/EEG experiment (see below), participants completed the  
50  
51 Mill Hill vocabulary test (Raven, Raven, & Court, 1998).  
52  
53  
54  
55  
56  
57  
58  
59  
60

Table 3 here

1 Sentence continuations from each participant were scored for whether or not they matched  
2 the critical resolved/unresolved sentence. We took only the first word from each response. These  
3 first-word responses were checked for spelling errors and corrected when the intended word was  
4 obvious (6 responses were excluded for being nonwords and therefore uninterpretable). We also  
5 checked whether the first-word response was a repetition of the final word(s) in the cut-off sentence  
6 and corrected where necessary (e.g. sentence: “The man asked about the nuggets and was told they  
7 were...”, response: “were chicken.”). Data sets from five participants were excluded (and replaced)  
8 because (1) they produced 9 or more (5+%) nonresponses or unusable/uninterpretable responses,  
9 and/or (2) they scored less than 33% correct on the vocabulary test (i.e. 2.5 SDs below the sample  
10 mean from the main MEG study). From the 11,520 trials (72 participants x 160 sentences), 47  
11 missing and uninterpretable responses were removed, resulting in 11,473 responses for inclusion in  
12 the analysis. A response was scored as a match if it was (1) an exact match, (2) an inflected form of  
13 the target word (e.g., “tastes” responses matched the target word “taste”), or (3) a longer or  
14 contracted form of the target word (e.g., “gymnasium” responses matched the target word “gym”).  
15 Responses were combined over participants, lead-in variants and versions.

16 For each of the 80 experimental items, we calculated the proportions of responses that  
17 matched the resolved sentence-final words (e.g. *tennis* and *game*) for sentences containing the  
18 ambiguous and control words (e.g. ACE and SPRINT). The resulting cloze probabilities for the  
19 critical words in our sentences were low overall (see Table 3; cloze probabilities for all stimuli can  
20 be downloaded from: <https://osf.io/3jhtb/>) confirming that – as intended – the sentence-final words  
21 were only weakly constrained by the preceding context. As the distributions of cloze probabilities  
22 for the four conditions were highly skewed, with high frequencies of 0 and near-0 cloze  
23 probabilities (i.e., cases where participants never or very rarely responded with the  
24 resolved/unresolved sentence-final word), we log-transformed the cloze probabilities to make these  
25 distributions more normal. Before this transformation, any probabilities of 0 were changed to a

lower-bound probability ( $\frac{1}{2}$  divided by the total number of responses for that condition), in order to avoid undefined values that result from taking the natural log of 0.

To quantify the degree of experimental control achieved in our materials, log-transformed cloze probabilities were entered into a Bayesian repeated-measures ANOVA with default priors (Morey, Rouder, & Jamil, 2015; Rouder, Morey, Speckman, & Province, 2012; Team, 2019). This analysis allows us to test for reliable differences in cloze probabilities between conditions as in a conventional ANOVA, but importantly to also assess evidence for the null hypothesis (i.e. that our sentence materials were well-matched as intended). We included within-item factors for word type (ambiguous or control) and sentence-final-word response type (resolved or unresolved word). Model comparisons provide very strong evidence for a difference between resolved and unresolved words ( $BF_{10}=43.217$ ) – indicating, as expected that, the more specific resolved words (e.g., *tennis*) were less predicted than the more generic unresolved words (e.g., *game*). Model comparisons provide moderate evidence for the null hypothesis that there is no difference between cloze probabilities following ambiguous (e.g., ACE) and control (e.g., SPRINT) words, ( $BF_{10}=0.130$ ). Most importantly, however, model comparisons also provide moderate evidence for the null hypothesis that the interaction between ambiguity and resolved/unresolved final words is absent ( $BF_{10}=0.258$ ). Based on standard interpretations of Bayes Factors (M. D. Lee & Wagenmakers, 2013) this suggests that it is approximately four times more likely that the interaction is absent than present. This therefore makes us confident that any interaction in MEG/EEG response amplitude at the sentence-final word, will not be attributable to differences in cloze probabilities.

### *Participants*

Twenty right-handed native British English speakers with normal hearing and no record of neurological diseases took part in the study for financial compensation. Ethical approval was issued by Cambridge Psychology Research Ethics Committee (University of Cambridge) and informed written consent was obtained from all volunteers. No participants had taken part in any of the pre-

1 tests described or had previously heard the sentences used. Data from four participants were  
2  
3  
4 excluded because of high noise in MEG or EEG (greater than 50% of trials were rejected during  
5  
6 data processing, see Methods); we report data from 16 participants (10 female), aged 20-39 years  
7  
8 (mean = 26.5, SD = 6 years).  
9

10  
11  
12  
13 *Experimental Procedure*  
14

15  
16 Experimental stimuli from list A or list B were presented auditorily (through in-ear headphones  
17  
18 connected via tubing to a pair of Etymotic drivers [www.etymotic.com](http://www.etymotic.com)) in four blocks (80 stimuli in  
19  
20 each block; 320 stimuli in total) interspersed with the fillers (20 stimuli in each block; 80 stimuli in  
21  
22 total) using EPrime 2 software (Psychology Software Tools). The four sentences from each stimulus  
23  
24 set appeared in separate blocks to avoid repetition of the key words within a block. Across  
25  
26 participants the order of blocks within the list was counterbalanced according to a Latin square  
27  
28 design such that each condition appeared before and after the other conditions an equal number of  
29  
30 times. Each participant heard a different pseudorandomised version for each block. Within a block  
31  
32 there were no more than three sequential presentations of an ambiguous stimulus and no more than  
33  
34 two sequential presentations of stimuli of a particular condition. There were no more than two  
35  
36 sequential presentations of fillers/task trials and no more than 10 trials between two fillers/task  
37  
38 trials.  
39  
40  
41

42  
43 Figures 1C and 1D show the structure of the experiment. The start of an experimental trial  
44  
45 was signalled to the listener by a red fixation cross (200 ms) visually presented on the screen,  
46  
47 during which they were encouraged to blink if necessary. The fixation turned black during a silent  
48  
49 period (jittered 1000±100ms) and remained on the screen throughout the duration of the spoken  
50  
51 sentence (2267 – 3765 ms) and for a post-sentence silent period (jittered 2000±100ms). The first  
52  
53 part of a filler/task trial followed an identical structure but spoken sentences were always followed  
54  
55 by a relatedness judgement task in which single words were presented visually (3000 ms) followed  
56  
57  
58  
59  
60



by a black fixation cross (jittered  $2000 \pm 100$  ms) and participants had to respond whether the word was related or unrelated to the meaning of the sentence they had just heard.

### *Behavioural measures*

Participants also performed a number of behavioural tasks allowing us to assess individual differences in comprehension skill, verbal knowledge, and non-verbal ability. Following the MEG/EEG recording we tested participants' comprehension of the critical sentences in which an ambiguous word was resolved to a subordinate meaning. Participants listened to the 80 ambiguous resolved sentences they had heard during the MEG/EEG session, each followed by auditory presentation of the ambiguous word from that sentence. They were asked to explain the meaning of that word as it was used in the preceding sentence, by typing in a synonym or a definition. They were not explicitly told that the words to which they had to respond were ambiguous. These responses were subsequently scored by a native-English speaker, naïve to the purpose of the experiment, who indicated whether participants generated the subordinate or dominant meaning of these words.

Participants' vocabulary knowledge was tested using the 34-question multiple-choice Mill Hill vocabulary test (Raven et al., 1998). We also measured participants' non-verbal ability with the Cattell 2a Culture Fair test (Cattell & Cattell, 1960), composed of four multiple choice subtests in which participants (1) complete a sequence of drawings, (2) select the odd one out from a set of drawings, (3) complete a pattern and (4) identify which drawing fulfils the criteria of an example. Following scoring of the individual behavioural tests, we assess across-participant correlations between test scores using Pearson correlations.

### *MEG and EEG data acquisition and pre-processing*

Magnetic fields were recorded (sampling rate 1000 Hz, bandpass filter 0.03-330 Hz) using a 306-channel Vectorview system (Elekta Neuromag, Helsinki) which contained one magnetometer and

two orthogonal gradiometer sensors at 102 locations within a helmet. Electric potentials were simultaneously recorded from 70 Ag/AgCl electrodes positioned according to the 10-10 system and embedded within an elasticated cap (Easy Cap). Additional electrodes positioned on the nose and one cheek were used as a reference and the ground respectively. Vertical and horizontal electro-oculograms were monitored with electrodes placed above and below the left eye, and either side of the eyes respectively. Electro-cardiogram was recorded with electrodes placed at the upper left and lower right area of the torso. Head position relative to the sensor array was recorded (using the Elekta Neuromag cHPI protocol with sampling rate of 200Hz) by using five head-position indicator (HPI) coils that emitted sinusoidal magnetic fields (293 – 321 Hz). Before the recording, the positions of the HPI coils and 70 EEG electrodes relative to three anatomical fiducials (nasion, left and right pre-auricular points) were digitally recorded using a 3D digitiser (Fastrak Polhemus). Approximately 80 additional head points over the scalp were also digitised to allow the offline reconstruction of the head model and co-registration with individual MRI images.

### *MEG and EEG data processing*

To minimise the contribution of magnetic sources from outside the head as well as any artifacts closer to the MEG sensor array, the data from the 306 MEG sensors were processed using the signal space separation method (Taulu & Kajola, 2005) and its temporal extension (Taulu & Simola, 2006), as implemented in Maxfilter 2.2 software (Elekta Neuromag): MEG sensors that generated poor quality data were identified and data interpolated, magnetic interference from non-neural sources was suppressed (tSSS buffer of 10 ms and correlation threshold of 0.98). Within-block movement in head position (as measured by HPI coils with HPI step set to 10ms) were compensated and data interpolated to adjust for head movement between blocks (interpolation to the first block). Finally data were downsampled to 250Hz.

Subsequent pre-processing was performed using MNE Python version 0.14 (Gramfort et al., 2013; Gramfort et al., 2014). For each participant, continuous data from the four recording blocks

were concatenated, visually inspected and bad EEG channels identified. To identify components associated with eye blinks and cardiac activity and reduce their contribution to the data, an Independent Component Analysis (ICA, fastICA method) was performed on the raw data (filtered 1-45Hz, data from bad EEG channels excluded). Prior to fitting and applying the ICA, the data were whitened (de-correlated and scaled to unit variance - “z-standardised” - also called sphering transformation) by means of a Principle Component Analysis (PCA). The number of PCA components entering the ICA decomposition was selected such that a cumulative variance of 0.9 was explained. Bad EEG channels were interpolated after ICA using spherical spline interpolation (Perrin, Pernier, Bertrand, & Echallier, 1989), continuous data filtered (4<sup>th</sup> order Butterworth, 0.1 to 40 Hz), and EEG data were re-referenced to the average over all EEG channels suitable for source analysis. Long epochs were created around the offset of the critical words at the two time points of interest (Ambiguity -2800 to 2500 ms; Disambiguation -4400 to 1500 ms) and each data point baseline-corrected using mean amplitude in the silent period before the sentence onset (Ambiguity -2800 to -2400 ms; Disambiguation -4400 to -4000 ms).

We chose to time-lock MEG and EEG responses to word offset because at this point listeners would have sufficient phonological information to recognise the critical words. Since many of our critical words were monosyllabic, word recognition was unlikely to occur before this time point (Marslen-Wilson, 1987). Subsequent processing and analyses were performed on shorter epochs before and after these word offsets (Ambiguity: -200 to 800 ms and Disambiguation: -500 to 1500 ms). These time windows were chosen in advance based on our expectations regarding the timing of neural responses associated with initial meaning selection and reinterpretation and on the known timing of the critical words in our stimuli (Figure 1B). In all sentences there was at least 800 ms between the ambiguous-word-offset and disambiguation-word offset (Figure 1B right panel, dotted line), and in 81% of sentences there was at least 800 ms between ambiguous-word-offset and disambiguation-word onset (Figure 1B right panel, solid line), thus we could be confident that effects before 800 ms should be attributable to initial meaning selection triggered by the ambiguity

1 rather than subsequent reinterpretation triggered by the disambiguating word. Epochs were rejected  
2  
3  
4 when peak-to-peak amplitudes within the epoch exceeded the following thresholds: 1000 fT/cm in  
5  
6  
7 gradiometers, 3500 fT in magnetometers, and  $120\mu\text{V}$  in EEG (mean rejection rates: targets 13.3%  
8  
9 trials, sentence-final words 21.1% trials) and the remaining epochs were averaged across  
10  
11 conditions.  
12

13  
14  
15  
16 *Sensor-space analysis*

17 Before analysis, between-participant differences in head positions within the helmet were calculated  
18  
19 and compensated. To do this we calculated the mean sensor array across participants then identified  
20  
21 the participant closest to this average (according to both translation and rotation parameters). MEG  
22  
23 data from all participants were transformed to this common sensory array using the ‘-trans’ option  
24  
25 in MaxFilter 2.2 software (Elekta Neuromag). Data were then analysed separately for gradiometers,  
26  
27 magnetometers and EEG. Before the gradiometer analysis, for every participant and condition, data  
28  
29 from each of the 102 sensor pairs were combined by taking the Root Mean Square (RMS) of the  
30  
31 two amplitudes:  $rms(g) = \sqrt{\frac{g1^2 + g2^2}{2}}$ . This is a standard procedure in MEG analysis, which  
32  
33  
34  
35  
36  
37 removes information about the direction of the two orthogonal gradients at each location. The  
38  
39 directions of the gradients vary across locations with respect to the brain and thus are not  
40  
41 meaningful for the purposes of our experimental questions. Before EEG analysis, the data were re-  
42  
43 referenced to the average of left and right mastoid recordings, to allow data to be more comparable  
44  
45 to most previous research on language (note that average referencing is required for  
46  
47 combined MEG/EEG source analysis).  
48  
49

50 Between-condition differences were assessed using non-parametric cluster-based  
51  
52 permutation tests (Maris & Oostenveld, 2007) to correct for multiple comparisons in time and  
53  
54 space. Using this method, conditions were compared and a t-value calculated for every time point  
55  
56 and every sensor. All samples with t-values greater than a threshold equivalent to  $p < .05$  ( $t = 1.753$ ,  
57  
58 one-tailed;  $t = 2.131$ , two-tailed) were selected and clustered based on temporal and spatial  
59  
60

adjacency, then cluster-level test statistics were calculated by summing all t-values in a cluster. To evaluate significance, the maximum cluster-level test statistic was compared against a null distribution generated by permutations: the subject-specific averages were randomly permuted within each subject (5000 times) and the Monte Carlo method used to create an approximation of the distribution of the test statistics under the null hypothesis. The Monte Carlo p-value is the proportion of cluster-level test statistics from the permutation distribution that are larger than the observed cluster-level test statistic. Clusters in which the p-value was smaller than the critical alpha-level of 0.05 support the conclusion that the two conditions are significantly different. Across participants, we tested for correlations between the amplitude of neural responses and behavioural scores, using the mean amplitude across the significant sensor-time points within the cluster.

Analyses focused on responses at the time of ambiguity and at the time of disambiguation (Figure 1A). To identify neural processes associated with initial meaning activation or selection at the time of ambiguity, we tested for a directional main effect of Ambiguity: i.e. whether ambiguous words elicit greater neural responses than the unambiguous control words. To identify neural processes associated with reinterpretation at the time of disambiguation, we tested for a directional interaction between Ambiguity and Disambiguation. The interaction allowed us to avoid confounds due to differences in the informativeness of the sentence-final words within each stimulus set (e.g. *tennis* necessarily has a more specific meaning than *game*). Specifically, disambiguating sentence-final words that resolve the ambiguity to a subordinate meaning should elicit greater activity than sentence-final words that leave the ambiguity unresolved and this difference in activation should be greater than the difference between responses to the acoustically identical sentence-final words in an unambiguous sentence. For the gradiometer analyses we performed one-tailed tests because the data had been rectified using RMS transformation and so values were all positive and monotonically linked to underlying neural activity. We could therefore be confident that ambiguous words would lead to increased signal compared to control words. For magnetometer and EEG analyses we performed two-tailed tests because we did not have specific predictions regarding the

1 polarity of sensor-level effects. Correlation analyses assessing individual differences in  
2 comprehension were all 2-tailed since even for comparisons in which we can be confident of  
3 observing greater activity for ambiguous than for control items (e.g. ambiguous vs control items for  
4 gradiometers), we could not anticipate whether more successful ambiguity resolution would be  
5 associated with greater neural engagement or reduced processing effort (see Taylor, Rastle, &  
6 Davis, 2013; 2014 for discussion).  
7  
8  
9  
10  
11  
12  
13  
14

15  
16  
17  
18 *Source Estimation*

19 To estimate the neural sources underpinning the observed sensor data, we used SPM 12 (Wellcome  
20 Trust Centre for Neuroimaging, London, UK). Data from all three neurophysiological measurement  
21 modalities (EEG and MEG magnetometers and gradiometers) were integrated using multimodal  
22 source inversion, which has been shown to give more precise localization than that obtained by  
23 considering each modality in isolation (Henson, Mouchlianitis, & Friston, 2009). With such an  
24 approach, sensor types with higher estimated levels of noise contribute less to the resulting source  
25 solutions.. For each participant, high-resolution structural MRI images (T1-weighted) were obtained  
26 using a GRAPPA 3D MPRAGE sequence (resolution time = 2250 ms; echo time = 2.99 ms; flip  
27 angle = 9% and acceleration factor = 2) on a 3T Tim Trio MR scanner (Siemens, Erlangen,  
28 Germany) with 1x1x1 mm isotropic voxels. For each individual, the structural MRI image was  
29 normalised to the standard Montreal Neurological Institute (MNI) template brain. The inverse  
30 normalisation parameters were then used to spatially transform canonical meshes for the cortex  
31 (8196 vertices), and scalp and skull (2562 vertices) to the individual space of each participant's  
32 MRI. Sensor locations and the scalp meshes were aligned using the 3 fiducial points measured  
33 during digitisation with those identified on the MRI scan, and with the digitised head shape.  
34  
35 Forward models to specify how any given source configuration appears at the sensors, were created  
36 separately for MEG using a single-shell model and for EEG using a boundary element model  
37 (following the recommendations specified in Litvak et al. (2011)).  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Source inversion was performed using the distributed minimum norm method (no depth weighting), which attempts to minimise overall source power whilst assuming all currents are equally likely to be active (Dale et al., 2000). An additional constraint was imposed (SPM "group inversion, as recommended in Litvak et al., 2011), whereby responses for all participants should be explained by the same set of sources, which has been shown to improve group-level statistical power (Litvak & Friston, 2008). In brief, the procedure involves 1) realigning and concatenating sensor-level data across subjects 2) estimating a single source solution for all subjects 3) using the resulting group solution as a Bayesian prior on individual subject inversions. Thus, this method exploits the availability of repeated measurements (from different subjects) to constrain source reconstruction. Importantly, however, the method does not bias activation differences between conditions to a given source. Source power (equivalent to the sum of squared amplitude) in the 0.1-40 Hz range was calculated from the resulting solutions and converted into 3D images. Significant effects from sensor space were localised by taking the mean 3D source power estimates across the relevant time windows and mapping the data onto MNI space brain templates. Between-condition differences were calculated and statistical significance in each voxel assessed with a series of one-sampled t-tests at the group level (i.e. mean signal divided by cross-participant variability). Since the aim of the source reconstruction was to localise significant sensor-space effects, results are displayed with an uncorrected voxel-wise threshold ( $p < .05$ , Gross et al., 2013)

## Results

### *Behavioural results*

On the semantic relatedness judgement task participants scored highly overall (mean proportion correct = 0.93, SD = 0.05) indicating they had listened attentively to the sentence stimuli. Overall, participants performed well on the post-MEG/EEG comprehension test indicating successful disambiguation of the ambiguous resolved sentences (mean = 0.94, SD = 0.04 proportion correct; scores for one participant were inadvertently not recorded resulting in  $n=15$  for analyses of

1 comprehension scores). Non-verbal IQ scores were above average for the general population (mean  
2 = 130.3, SD = 15.8 normalised scores). On average, participants knew around two thirds of the  
3 words in the vocabulary test (mean = 0.63, SD = 0.12 proportion correct). Correlational analysis  
4 revealed a positive correlation between sentence comprehension and vocabulary scores ( $r(15)=.638$ ,  
5  $p = .0105$ ; Figure 2). There were no reliable correlations between any of the other behavioural  
6 measures.

15 Figure 2 here

16 *MEG/EEG responses at the time of ambiguity*

17 Statistical analysis in sensor space revealed significant effects for gradiometers only (there were no  
18 significant clusters for magnetometers or EEG). At the offset of the ambiguous word there was  
19 significantly greater activity in response to ambiguous compared to unambiguous control words  
20 observed in a single sensor-time cluster from approximately 400 to 800 ms after word offset and  
21 most pronounced over left fronto-temporal sensors (cluster: 392 to 800 ms,  $p = .034$ , 1-tailed,  
22 Figure 3). Across participants, the amplitude of this response (averaged over significant sensor-time  
23 points) showed a marginally significant positive correlation with comprehension scores ( $r(14) =$   
24  $0.51$ ,  $p = .052$ , 2-tailed; Figure 4). These analyses included responses to all trials irrespective of  
25 whether the sentence was correctly interpreted. To further explore the relationship between MEG  
26 responses and successful comprehension we reanalysed the data excluding trials from sentences that  
27 were incorrectly understood in the post-MEG comprehension test (one participant was excluded due  
28 to a failure in recording the comprehension data). The MEG response at the time of ambiguity  
29 remained statistically reliable (cluster 372 to 800 ms,  $p = .025$ , 1-tailed) and the cross-participant  
30 correlation with comprehension remained marginally significant ( $r(14) = 0.45$ ,  $p = .092$ , 2-tailed).  
31 Since, on average, only 6% of sentences were misunderstood there were insufficient trial numbers  
32 to explore comprehension failures in more detail.

33 To confirm that the ambiguity response occurred prior to the presentation of any  
34 disambiguating information, we carried out a post-hoc analysis in which we excluded those trials in  
35



which the sentence for at least one condition had less than an 800 ms delay between target word offset and the onset of disambiguating words, i.e., those items for which our analysis window could include a response to the onset of disambiguation words. This resulted in the exclusion of the sentences for 19/80 ambiguous words (76/320 sentences per participant); in the remaining sentences the onset of the disambiguating word started after the end of the analysis time window (defined apriori as -200 to 800 ms relative to target word offset). Reanalysis of this subset of trials still showed a significant ambiguity effect (cluster 304 to 800 ms,  $p = 0.38$ , 1 tailed) confirming that these effects are due to ambiguous words and not subsequent disambiguation.

Figure 3 here

Figure 4 here

Source localisation of the significant neural response to ambiguous words showed cortical generators in fronto-temporal regions bilaterally (Figure 5; numbered source clusters are reported in Table 4). On the left, increased power for ambiguous compared to unambiguous control words was seen in the anterior portion of the ITG extending posteriorly and on the border with MTG (cluster 1, cluster 15). On the right, there was an area of activation in homologous regions of the ITG (cluster 6), which extended into the MTG (cluster 3), and a small cluster in superior temporal gyrus (STG, cluster 16). There was also a cluster in supramarginal gyrus (cluster 11). Frontally, there was a large right-lateralised cluster of activation in the IFG pars triangularis (cluster 7), extending into IFG pars orbitalis (cluster 8), and IFG pars opercularis (13), and in the middle frontal and superior frontal gyri (clusters 5 and 9). On the left, similar clusters of activation were seen in IFG pars opercularis (cluster 12) and middle frontal gyrus (cluster 10).

Table 4 here

Figure 5 here

#### *MEG/EEG responses at the time of disambiguation*

At the sentence-final word, non-parametric cluster-based permutation analysis revealed marginally significant interactions between Ambiguity and Disambiguation for gradiometers and for EEG.

1  
2 These arise from sensor-time clusters at around the time of word offset (Figure 6). For  
3  
4 gradiometers, the interaction corresponded to a cluster in the left and right hemisphere lasting from  
5  
6 approximately 200ms before to 200ms after the sentence-final word (cluster: -196 to 156 ms,  $p =$   
7  
8 .078, 1-tailed). For EEG, the interaction corresponded to a cluster for a central cluster of electrodes  
9  
10 from over a similar latency range (cluster: -276 to 212ms,  $p = .081$ , 2-tailed). As predicted, these  
11  
12 two marginal effects reflect greater activation for sentence-final words that resolved the ambiguity  
13  
14 to a subordinate meaning compared to words that left the ambiguity unresolved; no equivalent  
15  
16 difference was observed for resolved/unresolved words that completed the unambiguous sentences.  
17  
18 The EEG data also showed a marginally significant interaction for a sensor-time cluster in a later  
19  
20 time window (cluster 1144 to 1500 ms,  $p = .083$ , 2-tailed), but as can be seen in Figure 6C, the  
21  
22 effect was driven by a greater difference between sentence-final words in the unambiguous control  
23  
24 sentences than the ambiguous sentences. Since the direction of this interactional effect is  
25  
26 inconsistent with any specific functional contribution to reinterpretation we do not consider it  
27  
28 further.  
29  
30  
31  
32

33 To fully characterise the interaction of interest, we also performed post-hoc simple-effect  
34  
35 analyses. For the ambiguous sentences, sentence-final words that resolved the ambiguity elicited  
36  
37 greater activity than those which left the ambiguity unresolved corresponding to clusters in the  
38  
39 gradiometer (cluster -236 to 336 ms,  $p = 0.002$ , 1-tailed) and EEG data (cluster -196 to 236 ms,  $p =$   
40  
41 .047, 2-tailed). There was no significant effect for the unambiguous sentences (i.e. those that  
42  
43 contain a control word rather than an ambiguous word). There was also greater activation for words  
44  
45 resolving the ambiguity relative to acoustically identical words that completed an unambiguous  
46  
47 sentence (gradiometers: cluster -128 to 212 ms,  $p = .014$ , 1-tailed; cluster -436 to -100 ms,  $p = .059$ ,  
48  
49 1-tailed; EEG: cluster -172 to 226 ms,  $p = .028$ , 2-tailed) but no difference between the sentence-  
50  
51 final words that left the ambiguity unresolved compared to the same words that completed an  
52  
53 unambiguous sentence.  
54  
55  
56  
57  
58  
59  
60

Figure 6 here

To identify the source of the disambiguation effect, we performed source localisation on the time window -196 to 156 ms, covering the overlapping time period of effects in the MEG gradiometer and EEG analyses. We looked for regions with increased power for words that resolved the ambiguity than for words that left the ambiguity unresolved, compared to the equivalent difference in power between identical words that completed unambiguous sentences. Results (Figure 7, Table 5) show generators in left fronto-temporal regions including regions that overlap with those active at the time of ambiguity such as the ITG, extending to fusiform (cluster 2). There was also a cluster in IFG pars opercularis (cluster 6) and smaller frontal clusters in superior frontal gyrus (clusters 8 and 9), middle frontal gyrus (cluster 3), precentral gyrus (cluster 7), and supplementary motor area (cluster 13). On the right, there was a large cluster in supplementary motor area, extending to superior frontal gyrus and precentral gyrus (cluster 1), and in the middle frontal gyrus (cluster 4). We also saw bilateral clusters in the supramarginal gyrus (clusters 10 and 11).

Figure 7 here

Table 5 here

## Discussion

Using MEG/EEG we investigated the spatiotemporal dynamics of semantic ambiguity resolution by recording neural responses time-locked to the offset of an ambiguous word and to a subsequent disambiguating word that resolved the ambiguity to a subordinate meaning. Building on previous fMRI research, we capitalised on the high temporal resolution of MEG/EEG to distinguish between the neuro-cognitive processes of initial meaning access/selection versus reinterpretation. These are functionally distinct processes which in our sentences occur just a few hundred milliseconds apart. We feel confident that we have distinguished these neurocognitive effects for two reasons. First, an increased neural response associated with the processing of ambiguous words occurred before disambiguating information that triggers reinterpretation was presented. Second, neural manifestations of these processes were assessed with two orthogonal statistical contrasts: initial

1  
2 ambiguity processing was assessed through a main effect whereas reinterpretation was assessed  
3  
4 with an interaction.

5  
6 At the time of ambiguity, we observed significantly greater MEG responses for ambiguous  
7  
8 words versus unambiguous control words (Figure 3). The effect remained significant when we  
9  
10 excluded trials in which the onset of the sentence-final word that triggers reanalysis occurred within  
11  
12 the analysis window. Thus this neural effect of ambiguity was observed before the presentation or  
13  
14 processing of disambiguating information. Furthermore, the amplitude of the MEG response at the  
15  
16 time of ambiguity correlated positively with individual differences in comprehension skill, as  
17  
18 measured by our post-MEG comprehension test for ambiguous resolved sentences (Figure 4),  
19  
20 although this effect was only marginally significant. Comprehension also correlated positively with  
21  
22 vocabulary scores across participants (Figure 2). We discuss the cognitive processes associated with  
23  
24 these neural responses in the next section. In a subsequent section we then turn to neural responses  
25  
26 at the time of disambiguation; we observed marginally greater MEG and EEG response amplitudes  
27  
28 at the offset of sentence-final words that resolved an ambiguous word to a subordinate meaning  
29  
30 (Figure 6).  
31  
32  
33  
34

35  
36 Source estimation localised ambiguity responses to bilateral fronto-temporal regions (Figure  
37  
38 5) and disambiguation responses to bilateral frontal and left temporal regions (Figure 7). Given the  
39  
40 overlapping neural localisation of the two cognitively distinct processes involved in ambiguity  
41  
42 resolution, we will discuss these findings from source localisation in a final section of the  
43  
44 discussion, drawing on comparisons with the fMRI literature to inform our functional interpretation  
45  
46 of these neural responses.  
47  
48  
49  
50

51  
52 *Functional significance of neural responses to ambiguity*  
53

54  
55 We take the increased neural response after the offset of ambiguous words to reflect more  
56  
57 effortful processing of words with more than one meaning compared to matched single-meaning  
58  
59 control words. More specifically we relate the effect to the increased demands of meaning access  
60

1 and selection when multiple possible meanings are known. This neural effect is consistent with  
2 fMRI studies, as well as data from eye-tracking and ERP studies on the processing of visually-  
3 presented ambiguous words in sentences that we reviewed in the introduction.  
4  
5  
6  
7

8 While we described the observed response to ambiguity as a neural correlate of initial  
9 meaning activation or selection, which we distinguish from subsequent reinterpretation, this still  
10 leaves details of its functional contribution unspecified. It is, thus far, unclear whether the  
11 ambiguity response reflects processes involved in either (i) accessing and maintaining multiple  
12 meanings, or (ii) selecting a single meaning of an ambiguous word (e.g. by boosting or suppressing  
13 one or other meaning). Both these processes could be more engaged and/or more demanding for  
14 words with multiple meanings and hence plausibly observed in our comparison of responses to  
15 ambiguous and control words. Critical for distinguishing these two processes is the time-course  
16 over which listeners select a single meaning of an ambiguous word for sentences in which prior  
17 context does not constrain the likely meaning (as in the present experiment). However, conventional  
18 univariate analysis of MEG/EEG data cannot provide information on whether and when both  
19 meanings of an ambiguous word are active.  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34

35 Several sources of experimental evidence have been used to infer the time-course of  
36 meaning selection of ambiguous words in neutral context sentences. For example, cross-modal  
37 priming studies from Seidenberg et al. (1982) and Swinney (1979) are consistent with initial access  
38 to multiple meanings followed by selection of a single, dominant meaning. Swinney (1979)  
39 provides evidence for selective access by 3 syllables after word offset, in the time range of 750 –  
40 1000 ms (p.657), whereas Seidenberg et al. (1982) suggest it can occur sooner, within 200ms of  
41 word offset (both studies indicate activation of both meanings at word offset but do not test  
42 additional time points). Since both sets of data include a speeded response task with latencies  
43 between 500ms and 1000ms, of which around 150 ms can be accounted for in motor response  
44 planning, the minimum time-course over which multiple meanings are maintained before selection  
45 occurs is in the region of 550 - 950 ms. However, it is difficult to infer the specific timing of  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2 selection from these studies, in part because meaning activation is only measured indirectly (by  
3  
4 lexical decision or naming response times to targets related to one or other ambiguous word  
5  
6 meaning) at discrete points in time, and not to the ambiguous word itself. Nonetheless, in the  
7  
8 context of the present study, these findings would suggest that meaning selection takes place before  
9  
10 disambiguating information is presented for the majority of our sentences.  
11  
12

13         However, successful comprehension of most of our critical sentences ultimately depends on  
14  
15 selecting a lower frequency or subordinate meaning. Therefore initial selection of a single dominant  
16  
17 meaning (if that also entails full suppression of alternative meanings) would make reinterpretation  
18  
19 even more difficult. Yet, our post-MEG/EEG comprehension test showed that, on average, listeners  
20  
21 were able to understand more than 90% of delayed disambiguation sentences indicating that  
22  
23 reinterpretation was for the most part successful. Therefore, even if full suppression occurs listeners  
24  
25 can still semantically-reanalyse the sentence when they encounter a disambiguating word that  
26  
27 conflicts with the previously selected meaning (perhaps using phonological or working memory).  
28  
29 Alternatively, full suppression of alternative meanings may not occur and multiple meanings of  
30  
31 ambiguous words remain accessible and to some degree active, at least up to the point of  
32  
33 disambiguation. This proposal is consistent with response time data from a self-paced reading task  
34  
35 showing that multiple meanings can be maintained over even longer delays until disambiguation  
36  
37 (Miyake, Just, & Carpenter, 1992).  
38  
39  
40  
41  
42

43         One parsimonious description of longer-term maintenance of multiple meanings is through a  
44  
45 graded constraint-satisfaction process in which listeners make progressively stronger commitments  
46  
47 over time as evidence for alternatives increases (MacDonald, Pearlmutter, & Seidenberg, 1994). By  
48  
49 this account, neural activity after an ambiguous word reflects the activation of multiple alternative  
50  
51 interpretations in a representational space that also provides a mechanism for meaning maintenance  
52  
53 such that subsequent context can guide selection. In this account, there is therefore no separation of  
54  
55 the neural resources required for initial activation, maintenance in working memory, and meaning  
56  
57  
58  
59  
60

1  
2 selection. At face-value, this appears consistent with source localisation results that we discuss  
3  
4 below.

5  
6 One hallmark of this constraint-satisfaction account is that individual differences in sentence  
7  
8 comprehension arise from experience-dependent learning of the probabilities and regularities that  
9  
10 underlie language rather than in some external, capacity-limited system (such as working memory,  
11  
12 see MacDonald & Christiansen, 2002 for theoretical elaboration along with recurrent neural  
13  
14 network implementation). The present data provides tentative findings concerning the relationship  
15  
16 between individual differences in comprehension and neural responses to semantically ambiguous  
17  
18 words. For a sentence in which preceding context does not provide any specific information to  
19  
20 constrain word meaning activating and maintaining multiple semantic alternatives is optimal.  
21  
22 Hence, additional activation associated with ambiguous words should be associated with more  
23  
24 successful comprehension. In line with this proposal, we observed a positive correlation (albeit,  
25  
26 only marginally significant using a two-tailed test) between the amplitude of the ambiguity-related  
27  
28 MEG response and comprehension success in individual participants. The positive relationship  
29  
30 remained when we excluded sentences containing ambiguous words that specific participants did  
31  
32 not interpret correctly in the post-MEG/EEG comprehension test. This association is therefore not  
33  
34 explained by reduced responses to sentences for which listeners failed to correctly retrieve the  
35  
36 subordinate meaning. Thus, better comprehenders show greater neural processing effort in response  
37  
38 to ambiguous words.  
39  
40  
41  
42  
43  
44  
45

46 We explain this correlation between neural responses and comprehension as indicating that  
47  
48 successful comprehension of sentences containing ambiguous words requires additional processes  
49  
50 for activation and maintenance of alternative meanings. These result in increased availability of the  
51  
52 appropriate meaning which is required when subsequent context resolves the ambiguity to a  
53  
54 subordinate meaning. Interestingly, better comprehenders not only have increased availability of  
55  
56 subordinate meanings but also achieved higher vocabulary scores. It might be that higher-quality  
57  
58 lexical representations are required both for access to low-frequency meanings of unambiguous  
59  
60

1 words (for the more difficult items in the vocabulary test), and for accessing subordinate meanings  
2  
3 of ambiguous words (as in our MEG/EEG study). Nonetheless, given the small number of  
4  
5 participants and marginally significant results in the present study, this correlation between neural  
6  
7 activity and successful comprehension requires replication and extension. For example, we might  
8  
9 use more difficult sentences to directly compare neural activity associated with successful and  
10  
11 unsuccessful ambiguity resolution, or consider other predictors of individual variation to relate  
12  
13 ambiguity resolution (specifically) and spoken language comprehension (more generally). The  
14  
15 present study showed no association between non-verbal IQ and comprehension, but our  
16  
17 participants did not show as much variation in cognitive abilities as we might expect in the wider  
18  
19 population. More systematic exploration with a larger group of individuals with greater variability  
20  
21 in comprehension and measures of other cognitive factors (such as phonological short-term or  
22  
23 working memory) would be valuable.  
24  
25  
26  
27  
28  
29  
30  
31

32 *Functional interpretation of neural responses to disambiguation*

33 In addition to neural activity at the time of the ambiguous word we observed a potential neural  
34  
35 marker of reinterpretation during the presentation of sentence-final words that favour the  
36  
37 subordinate meaning of a previous ambiguous word. Importantly, reinterpretation effects observed  
38  
39 at sentence offset in both MEG and EEG were apparent as an interaction between the presence of an  
40  
41 ambiguous word and a sentence-final word that mandated access to an initially non-preferred,  
42  
43 subordinate meaning. This statistical interaction rules out the possibility that these effects are  
44  
45 responses simply to the presence of an ambiguous word, or a more informative sentence-final word  
46  
47 (the potentially-disambiguating words necessarily referred to more specific concepts and had a  
48  
49 lower cloze probability). Consistent with this conclusion, post-hoc simple effects showed that the  
50  
51 neural response to a sentence-final word was affected by the presence of an ambiguous word earlier  
52  
53 in a sentence only when the sentence-final word disambiguated the ambiguity (and not if the  
54  
55 sentence-final word left the ambiguity unresolved). Similarly, response differences between  
56  
57  
58  
59  
60



1  
2 ambiguous and control words were only apparent at sentence offset if the sentence-final word  
3  
4 served to resolve the ambiguity (but not if the sentence-final word did not conflict with the  
5  
6 dominant meaning of the ambiguous word). While the neural responses associated with  
7  
8 reinterpretation in MEG (gradiometers) and EEG were only marginally significant in analyses  
9  
10 correcting for time and sensors, the same pattern of neural difference was observed in both  
11  
12 modalities, and in overlapping time windows. This similarity gives us greater confidence in the  
13  
14 reliability of these observations.  
15  
16

17         The approximate timing and sensor topography of neural responses to reinterpretation are  
18  
19 broadly consistent with interpretation as an N400 effect (Kutas & Hillyard, 1980). Although the  
20  
21 N400 has been frequently observed in the EEG and MEG literature on language processing and  
22  
23 known to be associated with the processing of meaning, as yet there is no consensus on an  
24  
25 underlying functional account or computational mechanisms (for a review, see Kutas & Federmeier,  
26  
27 2011). For example, cognitive accounts suggest it may reflect the ease of accessing information in  
28  
29 semantic memory (Kutas & Federmeier, 2000) or of integrating semantic information into context  
30  
31 (Van Berkum, 2009). Computationally it may be more generally characterised as a semantic  
32  
33 prediction error signal (Rabovsky & McRae, 2014), linked to changes in a probabilistic  
34  
35 representation of sentence meaning (Rabovsky, Hansen, & McClelland, 2018).  
36  
37  
38  
39

40         ERP N400 responses have previously been observed in response to disambiguating words  
41  
42 that resolve an ambiguity to its subordinate meaning (Gunter et al., 2003; Hagoort & Brown, 1994),  
43  
44 although as discussed in the introduction, there are several differences between these previous  
45  
46 studies and ours. First, in previous work, sentences were visually presented word-by-word whereas  
47  
48 our sentences were presented auditorily as connected speech. Secondly, previous studies did not  
49  
50 control for both the presence/absence of ambiguity and the word form itself. We showed a statistical  
51  
52 interaction between these two factors for sentence final words that trigger reinterpretation effects.  
53  
54 Unlike previous studies this interaction cannot be due to simple differences in word form or  
55  
56 meaning between the critical words in our sentences.  
57  
58  
59  
60

1  
2 One possibility raised by a reviewer was that the neural interaction generating this N400-  
3  
4 like response to reinterpretation could arise from differences in cloze probability between sentence-  
5  
6 final words in our critical conditions. However, a sentence completion test on our materials showed  
7  
8 that cloze probabilities were low overall (the median cloze probability was zero in both conditions  
9  
10 that contained the resolved word, close to zero for the unresolved words, and did not differ between  
11  
12 ambiguous and control words). We did not include highly constrained sentences, or semantic  
13  
14 anomalies that are typical of N400 studies. More importantly, though, a Bayesian analysis of cloze-  
15  
16 probability values provided moderately strong evidence that there was not an interaction between  
17  
18 ambiguity and reinterpretation (i.e. this analysis provides evidence that the sentences in our critical  
19  
20 conditions were matched for cloze probability). Hence, we can conclude that our N400-like effect  
21  
22 of reinterpretation is not due to variation in the ease of meaning access due to cloze probability, but  
23  
24 rather due to sentence-final words triggering reinterpretation. Nonetheless, future work to determine  
25  
26 the functional nature of the neural response to reinterpretation would benefit from comparing this  
27  
28 response to the semantic error response evoked by a sentence final anomalous word. Anomalous  
29  
30 words should trigger an N400-like response but would not result in reinterpretation and hence  
31  
32 differences between anomalous words and words driving reinterpretation may be informative.  
33  
34  
35  
36  
37  
38  
39  
40

41 *The role of fronto-temporal regions in ambiguity resolution*

42  
43 With regard to the anatomical questions that motivated the present study, our source  
44  
45 localisation provides evidence that frontal and temporal lobe regions are activated both in response  
46  
47 to ambiguous words in a neutral context (prior to presentation of disambiguating information;  
48  
49 Figure 5), and subsequently in response to a disambiguating word which resolves the ambiguity to a  
50  
51 subordinate meaning (Figure 7). Previous fMRI evidence has similarly demonstrated the  
52  
53 involvement of fronto-temporal regions in ambiguity resolution (Mason & Just, 2007; Musz &  
54  
55 Thompson-Schill, 2017; Rodd et al., 2005; Rodd et al., 2012; Rodd, Longe, et al., 2010; Vitello et  
56  
57 al., 2014; Zemleni et al., 2007). However, unlike in fMRI, timing information from MEG/EEG  
58  
59  
60

allows us to confidently attribute our ambiguity and disambiguation responses specifically to initial processing of the ambiguous word and also to subsequent reinterpretation of the ambiguous word. Initial meaning activation/selection of an ambiguous word was identified through a statistical main effect whereas subsequent reinterpretation at a disambiguating word was identified through a statistical interaction. Furthermore, responses associated with initial meaning activation/selection and subsequent reinterpretation could be separated in time; the neural response to ambiguity occurred before the onset of disambiguating words that trigger reinterpretation. Thus, these are two independent effects and overlap of the neural sources can inform our understanding of the underlying mechanisms.

As we reviewed in the introduction, previous fMRI studies on ambiguity resolution have associated activation in IFG regions on the left (Vitello et al., 2014) or bilaterally (Mason & Just, 2007; Rodd et al., 2012; Zempleni et al., 2007) with reinterpretation, and in one study, activation extended into superior and middle frontal areas (Mason & Just, 2007) in line with the left IFG and superior and middle frontal clusters shown here. Only two previous fMRI studies on ambiguity resolution tentatively associated initial meaning selection with activation in IFG (Mason & Just, 2007; Rodd et al., 2012). Consistent with previous conclusions and our findings that IFG is active both during initial meaning selection and subsequent reinterpretation, one dominant proposal regarding the functional role of the left IFG is its involvement in selecting between competing semantic representation (Jefferies, 2013; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997), or resolving conflict arising from competing stimulus representations of any format (Novick, Trueswell, & Thompson-Schill, 2005; for the suggestion that IFG activation is involved in selection (or conflict resolution) rather than simply reflecting increased competition between semantic representations, see Grindrod et al. (2008)).

An alternative account of IFG contributions to language, the Unification account (Hagoort, 2005, 2013), proposes a more general role for the IFG in combining individual words into coherent sentence- and discourse-level representations. These are processes which we might also expect to be

1  
2 taxed as the number of meanings increases and multiple meanings are accessed, maintained, or  
3  
4 predicted. While we cannot offer any evidence to adjudicate between these views, we argued above  
5  
6 that meaning selection of the ambiguous words in our study is likely not completed during the time  
7  
8 window before disambiguation. This seems to favour a more graded rather than absolute form of  
9  
10 selection, perhaps consistent with a constraint satisfaction or unification account.  
11  
12

13       Previous fMRI studies on ambiguity resolution associated activation in left MTG and ITG/  
14  
15 fusiform to reinterpretation (Rodd et al., 2005; Rodd et al., 2012; Rodd, Longe, et al., 2010; Vitello  
16  
17 et al., 2014; Zempleni et al., 2007). In line with this, localisation of the MEG/EEG response to a  
18  
19 disambiguating word indicated a source in left ITG and fusiform, which we attribute to  
20  
21 reinterpretation. Notably, we also observed neural sources of the MEG response to an ambiguous  
22  
23 word in MTG and ITG bilaterally, which could be linked to initial meaning activation or selection.  
24  
25 Posterior temporal regions have often been proposed to contribute to meaning access for isolated  
26  
27 words (see Hickok & Poeppel, 2007; Lau, Phillips, & Poeppel, 2008). These regions would  
28  
29 plausibly show greater activation when listeners access multiple meanings of ambiguous words:  
30  
31 first when ambiguity is initially encountered and again at a disambiguating word inconsistent with  
32  
33 the previously preferred meaning which triggers an increase in activation of an alternative. We also  
34  
35 note that left posterior MTG activation has previously been observed in response to syntactically  
36  
37 ambiguous words, using fMRI (Snijders et al., 2009) and MEG (Tyler, Cheung, Devereux, &  
38  
39 Clarke, 2013), although a recent meta-analysis suggests that these posterior temporal regions are  
40  
41 recruited more for semantic rather than syntactic processing (Rodd, Vitello, Woollams, & Adank,  
42  
43 2015).  
44  
45  
46  
47  
48

49       We earlier characterised the MEG/EEG reinterpretation effect as resembling an N400. In  
50  
51 line with this proposal we note there is some overlap between source localisation of the  
52  
53 reinterpretation response to the left ITG and IFG, and regions proposed to underpin the classic  
54  
55 N400 effect, which have been explored using fMRI and MEG/EEG (Halgren et al., 2002; Lau,  
56  
57 Gramfort, Hämäläinen, & Kuperberg, 2013; Lau, Weber, Gramfort, Hamalainen, & Kuperberg,  
58  
59  
60

2016; Maess, Herrmann, Hahne, Nakamura, & Friederici, 2006). The N400 is likely to reflect a combination of neural processes originating in multiple cortical regions sources but across a number of studies, it has been proposed that the effect may originate in posterior temporal regions before being observed in more anterior portions of the temporal lobe and IFG (for a review, see Lau et al., 2008).

Interestingly, the ITG/fusiform activations we observed at the time of disambiguation and in response to ambiguity extended to more anterior and inferior temporal regions than has been seen in previous fMRI studies of ambiguity resolution. Anterior temporal activations have been less consistently observed in fMRI, perhaps because standard EPI acquisitions give relatively poor signal in these regions (Devlin et al., 2000; Visser, Jefferies, & Lambon Ralph, 2010, although see Musz, 2017 for evidence of anterior inferior temporal representations of ambiguous words shown by MVPA fMRI). However, damage to the anterior temporal lobe has long been associated with impaired semantic processing in general (Patterson et al., 2007) and of semantically ambiguous words in particular (e.g. measured by patients' ability to produce alternative interpretations of unresolved ambiguous sentences (Zaidel et al., 1995). Thus, the inferior temporal activation we observed when listeners initially encounter an ambiguous word and when disambiguating information is heard is largely consistent with other evidence for semantic contributions of these basal temporal regions.

One other point to consider is that both the frontal and temporal neural sources of responses to ambiguity and disambiguation appear to be somewhat bilateral. Previous fMRI studies have reported significant activation in right frontal regions (Mason & Just, 2007; Rodd et al., 2005; Zempleni et al., 2007), although reports of right temporal lobe responses are more limited, and a meta-analysis of fMRI studies of semantic and syntactic processing demands reveals fewer and less reliable findings of right than left fronto-temporal activity (Rodd et al., 2015). However, in the absence of statistical comparison of left and right sided activity in fMRI or MEG/EEG we hesitate to draw strong conclusions from these observations (see, Peelle, 2012 for arguments that lateralised

1 effects in thresholded statistical maps provide little or no evidence for functional lateralisation).  
2  
3 Furthermore, other evidence is consistent with bilateral contributions to ambiguity resolution, for  
4  
5 example, from behavioural studies using lateralised word presentations (Burgess & Simpson, 1988;  
6  
7 Faust & Gernsbacher, 1996), and neuropsychological studies (Hagoort, 1993; Swaab, Brown, &  
8  
9 Hagoort, 1998; Tompkins, Baumgaertner, Lehman, & Fassbinder, 2000). While functional imaging  
10  
11 evidence can potentially play an important role in determining the differential contributions of the  
12  
13 left and right hemisphere to ambiguity resolution, published studies, including the present work,  
14  
15 have yet to report hemispheric dissociations sufficient to conclude that the left and right  
16  
17 hemispheres make distinct functional contributions to initial meaning activation and selection.  
18  
19  
20  
21  
22  
23  
24

25 *Conclusions*

26  
27 Taken together with previous fMRI research, our observations suggest that both temporal  
28  
29 and frontal regions play an important role both in initial meaning activation and selection for  
30  
31 ambiguous words, as well as later reinterpretation triggered by a disambiguating word. Previous  
32  
33 research has tried to fractionate frontal and temporal regions based on the time-course of activation  
34  
35 during delayed disambiguation sentences (Rodd et al, 2012) or by comparing responses to  
36  
37 ambiguous words with balanced and biased meaning frequencies (Mason & Just, 2007; Vitello et  
38  
39 al., 2014; Zempleni et al., 2007). However, source localisation results from MEG/EEG suggest that  
40  
41 frontal and temporal regions play a coordinated role both in the initial interpretation of ambiguous  
42  
43 words presented in neutral sentence contexts and subsequently when interpretations need to be  
44  
45 revised. This proposal could be taken to challenge traditional divisions between a temporal lobe  
46  
47 contributions to semantic representation and frontal contributions to working memory or selection  
48  
49 (see Musz & Thompson-Schill, 2017, for a recent statement along these lines).  
50  
51  
52  
53  
54

55 Rather than the traditional fractionation of temporal and frontal responses, we instead  
56  
57 propose a graded, constraint-satisfaction account which elides a simple distinction between  
58  
59 semantic representations and processing. In this account, neural activity after an ambiguous word  
60

reflects the activation of multiple alternative interpretations in a representational space that also supports neural mechanisms for meaning maintenance and eventual selection. During this time period selection can be construed as stronger, but not exclusive, activation of a particular meaning, which can only be confirmed when disambiguating information is presented. At this point, successful meaning integration and interpretation may require reinterpretation, which can be realised in terms of a re-weighting of the activation levels of different meanings. Future work to assess the representational dynamics of these frontal and temporal responses (e.g. using representational similarity or other, similar multivariate methods, Kriegeskorte, 2008) might provide additional evidence for this account.

## Funding

This work was supported by the Medical Research Council (SUAG/008 RG91365 to MHD)

## Acknowledgements

The authors thank Jane Warren for her help in creating the stimuli, Maarten van Casteren for helping during data collection, and Isobel Davis for help with scoring and analysing behavioural data.

- 1 Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database (CD-ROM)*.  
2 Philadelphia, PA.  
3  
4  
5  
6 Boynton, G. M., Engel, S. A., Glover, G. H., & Heeger, D. J. (1996). Linear systems analysis of  
7 functional magnetic resonance imaging in human V1. *Journal of Neuroscience*, 16(13),  
8 4207-4221.  
9  
10  
11  
12  
13 Burgess, C., & Simpson, G. B. (1988). Cerebral hemispheric mechanisms in the retrieval of  
14 ambiguous word meanings. *Brain and Language*, 33, 86-103.  
15  
16  
17 Cai, Z. G., Gilbert, R. A., Davis, M. H., Gaskell, M. G., Farrar, L., Adler, S., et al. (2017). Accent  
18 modulates access to word meaning: Evidence for a speaker-model account of spoken word  
19 recognition. *Cogn Psychol*, 98, 73-101.  
20  
21  
22  
23  
24  
25 Cattell, R. B., & Cattell, A. K. S. (1960). *Handbook for the individual or group Culture Fair*  
26 *Intelligence test*. Champaign, IL: Testing Inc.  
27  
28  
29 Coleman, M. R., Davis, M. H., Rodd, J. M., Robson, T., Ali, A., Owen, A. M., et al. (2009).  
30 Towards the routine use of brain imaging to aid the clinical diagnosis of disorders of  
31 consciousness. *Brain*, 132, 2541-2552.  
32  
33  
34  
35  
36 Coleman, M. R., Rodd, J. M., Davis, M. H., Johnsrude, I. S., Menon, D. K., Pickard, J. D., et al.  
37 (2007). Do vegetative patients retain aspects of language comprehension? Evidence from  
38 fMRI. *Brain*, 130(Pt 10), 2494-2507.  
39  
40  
41  
42  
43 Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., et al. (2000).  
44 Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution  
45 imaging of cortical activity. *Neuron*, 26(1), 55-67.  
46  
47  
48  
49  
50 Davis, M. H., Coleman, M. R., Absalom, A. R., Rodd, J. M., Johnsrude, I. S., Matta, B. F., et al.  
51 (2007). Dissociating speech perception and comprehension at reduced levels of awareness.  
52 *Proc Natl Acad Sci U S A*, 104(41), 16032-16037.  
53  
54  
55  
56  
57 de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web  
58 browser. *Behavior Research Methods*, 47(1), 1-12.  
59  
60



- Devlin, J. T., Russell, R. P., Davis, M. H., Price, C. J., Wilson, J., Moss, H. E., et al. (2000). Susceptibility-induced loss of signal: comparing PET and fMRI on a semantic task. *Neuroimage*, 11(6 Pt 1), 589-600.
- Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical Ambiguity and Fixation Times in Reading. *Journal of Memory and Language*, 27(4), 429-446.
- Faust, M. E., & Gernsbacher, M. A. (1996). Cerebral mechanisms for suppression of inappropriate information during sentence comprehension. *Brain Lang*, 53(2), 234-259.
- Federmeier, K. D., Segal, J. B., Lombrozo, T., & Kutas, M. (2000). Brain responses to nouns, verbs and class-ambiguous words in context. *Brain*, 123 Pt 12, 2552-2566.
- Frazier, L., & Rayner, K. (1990). Taking on semantic commitments: processing multiple meanings vs. multiple senses. *Journal of Memory and Language*, 29, 181-200.
- Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *J Exp Psychol Learn Mem Cogn*, 16(3), 430-445.
- Gilbert, R. A., Betts, H. N., Jose, R., & Rodd, J. M. (2017). *New UK-based dominance norms for ambiguous words*. Paper presented at the Poster session presented at the Experimental Psychology Society Meeting.
- Gramfort, A., Luessi, M., Larson, E., Engemann, D., Strohmeier, D., Brodbeck, C., et al. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, 7(267).
- Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., et al. (2014). MNE software for processing MEG and EEG data. *Neuroimage*, 86, 446-460.
- Grindrod, C. M., Bilenko, N. Y., Myers, E. B., & Blumstein, S. E. (2008). The role of the left inferior frontal gyrus in implicit semantic competition and selection: An event-related fMRI study. *Brain Res*, 1229, 167-178.
- Gross, J., Baillet, S., Barnes, G. R., Henson, R. N., Hillebrand, A., Jensen, O., et al. (2013). Good practice for conducting and reporting MEG research. *Neuroimage*, 65, 349-363.

- Gunter, T. C., Wagner, S., & Friederici, A. D. (2003). Working memory and lexical ambiguity resolution as revealed by ERPs: A difficult case for activation theories. *J Cogn Neurosci*, 15(5), 643-657.
- Hagoort, P. (1993). Impairments of Lexical Semantic Processing in Aphasia - Evidence from the Processing of Lexical Ambiguities. *Brain Lang*, 45(2), 189-232.
- Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends Cogn Sci*, 9(9), 416-423.
- Hagoort, P. (2013). MUC (Memory, Unification, Control) and beyond. *Front Psychol*, 4.
- Hagoort, P., & Brown, C. (1994). Brain Responses to Lexical Ambiguity Resolution and Parsing. *Perspectives on Sentence Processing*, 45-80.
- Halgren, E., Dhond, R. P., Christensen, N., Van Petten, C., Marinkovic, K., Lewine, J. D., et al. (2002). N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. *Neuroimage*, 17(3), 1101-1116.
- Henderson, L., Snowling, M., & Clarke, P. (2013). Accessing, Integrating, and Inhibiting Word Meaning in Poor Comprehenders. *Scientific Studies of Reading*, 17(3), 177-198.
- Henson, R. N., Mouchlianitis, E., & Friston, K. J. (2009). MEG and EEG data fusion: simultaneous localisation of face-evoked responses. *Neuroimage*, 47(2), 581-589.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nat Rev Neurosci*, 8(5), 393-402.
- Jefferies, E. (2013). The neural basis of semantic cognition: converging evidence from neuropsychology, neuroimaging and TMS. *Cortex*, 49(3), 611-625.
- Josephs, O., & Henson, R. N. A. (1999). Event-related functional magnetic resonance imaging: modelling, inference and optimization. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 354(1387), 1215-1228.
- Kambe, G., Rayner, K., & Duffy, S. A. (2001). Global context effects on processing lexically ambiguous words: evidence from eye fixations. *Mem Cognit*, 29(2), 363-372.

- 1  
2 Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in  
3  
4 language comprehension. *Trends Cogn Sci*, 4(12), 463-470.  
5  
6 Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400  
7  
8 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 62(1),  
9  
10 621-647.  
11  
12 Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic  
13  
14 incongruity. *Science*, 207(4427), 203-205.  
15  
16 Lange, K., Kuhn, S., & Filevich, E. (2015). "Just Another Tool for Online Studies" (JATOS): An  
17  
18 Easy Solution for Setup and Management of Web Servers Supporting Online Studies. *PLoS*  
19  
20 *One*, 10(6), e0130834.  
21  
22 Lau, E. F., Gramfort, A., Hämäläinen, M. S., & Kuperberg, G. R. (2013). Automatic Semantic  
23  
24 Facilitation in Anterior Temporal Cortex Revealed through Multimodal Neuroimaging. *The*  
25  
26 *Journal of Neuroscience*, 33(43), 17174-17181.  
27  
28 Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing  
29  
30 the N400. *Nat Rev Neurosci*, 9(12), 920-933.  
31  
32 Lau, E. F., Weber, K., Gramfort, A., Hamalainen, M. S., & Kuperberg, G. R. (2016).  
33  
34 Spatiotemporal Signatures of Lexical-Semantic Prediction. *Cereb Cortex*, 26(4), 1377-1387.  
35  
36 Lee, C. L., & Federmeier, K. D. (2006). To mind the mind: an event-related potential study of word  
37  
38 class and semantic ambiguity. *Brain Res*, 1081(1), 191-202.  
39  
40 Lee, C. L., & Federmeier, K. D. (2009). Wave-ering: An ERP study of syntactic and semantic  
41  
42 context effects on ambiguity resolution for noun/verb homographs. *J Mem Lang*, 61(4), 538-  
43  
44 555.  
45  
46 Lee, C. L., & Federmeier, K. D. (2012). Ambiguity's aftermath: how age differences in resolving  
47  
48 lexical ambiguity affect subsequent comprehension. *Neuropsychologia*, 50(5), 869-879.  
49  
50 Lee, M. D., & Wagenmakers, E. J. (2013). *Bayesian Modeling for Cognitive Science: A Practical*  
51  
52 *Course*: Cambridge University Press.  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2 Litvak, V., & Friston, K. (2008). Electromagnetic source reconstruction for group studies.  
3  
4 *Neuroimage*, 42(4), 1490-1498.
- 5  
6 Litvak, V., Mattout, J., Kiebel, S., Phillips, C., Henson, R., Kilner, J., et al. (2011). EEG and MEG  
7  
8 data analysis in SPM8. *Comput Intell Neurosci*, 2011, 852961.
- 9  
10 MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just  
11  
12 and Carpenter (1992) and Waters and Caplan (1996). *Psychol Rev*, 109(1), 35-54.
- 13  
14 MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic  
15  
16 ambiguity resolution [corrected]. *Psychol Rev*, 101(4), 676-703.
- 17  
18 Maess, B., Herrmann, C. S., Hahne, A., Nakamura, A., & Friederici, A. D. (2006). Localizing the  
19  
20 distributed language network responsible for the N400 measured by MEG during auditory  
21  
22 sentence processing. *Brain Res*, 1096(1), 163-172.
- 23  
24 Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *J*  
25  
26 *Neurosci Methods*, 164(1), 177-190.
- 27  
28 Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word recognition. *Cognition*,  
29  
30 25(1-2), 71-102.
- 31  
32 Mason, R. A., & Just, M. A. (2007). Lexical ambiguity in sentence comprehension. *Brain Res*,  
33  
34 1146, 115-127.
- 35  
36 Miyake, A., Just, M. A., & Carpenter, P. A. (1992). Working Memory Constraints on the  
37  
38 Maintenance of Multiple Interpretations of Lexical Ambiguities. *Bulletin of the*  
39  
40 *Psychonomic Society*, 30(6), 482-&.
- 41  
42 Morey, R. D., Rouder, J. N., & Jamil, T. (2015). BayesFactor: Computation of Bayes factors for  
43  
44 common designs.
- 45  
46 Musz, E., & Thompson-Schill, S. L. (2017). Tracking competition and cognitive control during  
47  
48 language comprehension with multi-voxel pattern analysis. *Brain Lang*, 165, 21-32.
- 49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

- Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive Affective & Behavioral Neuroscience*, 5(3), 263-281.
- Palan, S., & Schitter, C. (2017). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*.
- Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat Rev Neurosci*, 8(12), 976-987.
- Peelle, J. E. (2012). The hemispheric lateralization of speech processing depends on what "speech" is: a hierarchical perspective. *Front Hum Neurosci*, 6, 309.
- Peer, E., Samat, S., Brandimarte, L., & Acquisti, A. (2015). Beyond the Turk: An Empirical Comparison of Alternative Platforms for Online Behavioral Research. *SSRN Electronic Journal*.
- Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical Splines for Scalp Potential and Current-Density Mapping. *Electroencephalogr Clin Neurophysiol Suppl*, 72(2), 184-187.
- Rabovsky, M., Hansen, S. S., & McClelland, J. L. (2018). Modelling the N400 brain potential as change in a probabilistic representation of meaning. *Nature Human Behaviour*, 2(9), 693-705.
- Rabovsky, M., & McRae, K. (2014). Simulating the N400 ERP component as semantic network error: Insights from a feature-based connectionist attractor model of word meaning. *Cognition*, 132(1), 68-89.
- Raven, J., Raven, J. C., & Court, J. H. (1998). The Mill Hill vocabulary scale. In H. Assessment (Ed.), *Manual for Raven's progressive matrices and vocabulary scales*. San Antonio, TX.
- Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: effects of word frequency, verb complexity, and lexical ambiguity. *Mem Cognit*, 14(3), 191-201.

- Rodd, J. M., Cutrin, B. L., Kirsch, H., Millar, A., & Davis, M. H. (2013). Long-term priming of the meanings of ambiguous words. *Journal of Memory and Language*, 68(2), 180-198.
- Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cereb Cortex*, 15(8), 1261-1269.
- Rodd, J. M., Gaskell, G., & Marslen-Wilson, W. D. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language*, 46(2), 245-266.
- Rodd, J. M., Johnsrude, I. S., & Davis, M. H. (2010). The role of domain-general frontal systems in language comprehension: evidence from dual-task interference and semantic ambiguity. *Brain Lang*, 115(3), 182-188.
- Rodd, J. M., Johnsrude, I. S., & Davis, M. H. (2012). Dissociating frontotemporal contributions to semantic ambiguity resolution in spoken sentences. *Cereb Cortex*, 22(8), 1761-1773.
- Rodd, J. M., Longe, O. A., Randall, B., & Tyler, L. K. (2010). The functional organisation of the fronto-temporal language system: evidence from syntactic and semantic ambiguity. *Neuropsychologia*, 48(5), 1324-1335.
- Rodd, J. M., Vitello, S., Woollams, A. M., & Adank, P. (2015). Localising semantic and syntactic processing in spoken and written language comprehension: an Activation Likelihood Estimation meta-analysis. *Brain Lang*, 141, 89-102.
- Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, 56(5), 356-374.
- Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: some limitations of knowledge-based processing. *Cognitive Psychology*, 14, 489-537.
- Snijders, T. M., Vosse, T., Kempen, G., Van Berkum, J. J., Petersson, K. M., & Hagoort, P. (2009). Retrieval and unification of syntactic structure in sentence comprehension: an fMRI study using word-category ambiguity. *Cereb Cortex*, 19(7), 1493-1503.

- 1  
2 Swaab, T. Y., Brown, C., & Hagoort, P. (1998). Understanding ambiguous words in sentence  
3 contexts: electrophysiological evidence for delayed contextual selection in Broca's aphasia.  
4  
5  
6 *Neuropsychologia*, 36(8), 737-761.  
7
- 8  
9 Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of  
10 context effects. *Journal of Verbal Learning and Verbal Behavior*, 18, 645-659.  
11  
12  
13 Szabo Wankoff, L., & Cairns, H. S. (2009). Why Ambiguity Detection Is a Predictor of Early  
14 Reading Skill. *Communication Disorders Quarterly*, 30(3), 183-192.  
15  
16  
17 Tahmasebi, A. M., Davis, M. H., Wild, C. J., Rodd, J. M., Hakyemez, H., Abolmaesumi, P., et al.  
18  
19 (2012). Is the link between anatomical structure and function equally strong at all cognitive  
20 levels of processing? *Cereb Cortex*, 22(7), 1593-1603.  
21  
22  
23 Taulu, S., & Kajola, M. (2005). Presentation of electromagnetic multichannel data: The signal space  
24 separation method. *Journal of Applied Physics*, 97(12).  
25  
26  
27 Taulu, S., & Simola, J. (2006). Spatiotemporal signal space separation method for rejecting nearby  
28 interference in MEG measurements. *Phys Med Biol*, 51(7), 1759-1768.  
29  
30  
31 Taylor, J. S. H., Rastle, K., & Davis, M. H. (2013). Can cognitive models explain brain activation  
32 during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. 139(4),  
33  
34  
35 766-791.  
36  
37  
38 Taylor, J. S. H., Rastle, K., & Davis, M. H. (2014). Interpreting response time effects in functional  
39 imaging studies. *Neuroimage*, 99, 419-433.  
40  
41  
42 Team, J. (2019). JASP.  
43  
44  
45 Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior  
46 prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proc Natl Acad Sci U*  
47  
48  
49 *S A*, 94(26), 14792-14797.  
50  
51  
52  
53 Tompkins, C. A., Baumgaertner, A., Lehman, M. T., & Fassbinder, W. (2000). Mechanisms of  
54 discourse comprehension impairment after right hemisphere brain damage: Suppression in  
55  
56  
57  
58  
59  
60

lexical ambiguity resolution. *Journal of Speech Language and Hearing Research*, 43(1), 62-78.

Twilley, L. C., Dixon, P., Taylor, D., & Clark, K. (1994). University-of-Alberta Norms of Relative Meaning Frequency for 566 Homographs. *Memory & Cognition*, 22(1), 111-126.

Tyler, L. K., Cheung, T. P., Devereux, B. J., & Clarke, A. (2013). Syntactic computations in the language network: characterizing dynamic network properties using representational similarity analysis. *Front Psychol*, 4, 271.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, 15(1), 273-289.

Van Berkum, J. J. (2009). The neuropragmatics of 'simple' utterance comprehension: An ERP review. In K. Y. U. Sauerland (Ed.), *Semantics and pragmatics: From experiment to theory* (pp. 276-316). Basingstoke: Palgrave Macmillan.

Visser, M., Jefferies, E., & Lambon Ralph, M. A. (2010). Semantic processing in the anterior temporal lobes: a meta-analysis of the functional neuroimaging literature. *J Cogn Neurosci*, 22(6), 1083-1094.

Vitello, S., & Rodd, J. M. (2015). Resolving Semantic Ambiguities in Sentences: Cognitive Processes and Brain Mechanisms. *Lang Linguist Compass*, 9(10), 391-405.

Vitello, S., Warren, J. E., Devlin, J. T., & Rodd, J. M. (2014). Roles of frontal and temporal regions in reinterpreting semantically ambiguous sentences. *Front Hum Neurosci*, 8.

Zaidel, D. W., Zaidel, E., Oxbury, S. M., & Oxbury, J. M. (1995). The interpretation of sentence ambiguity in patients with unilateral focal brain surgery. *Brain Lang*, 51(3), 458-468.

Zempleni, M. Z., Renken, R., Hoeks, J. C. J., Hoogduin, J. M., & Stowe, L. A. (2007). Semantic ambiguity processing in sentence context: Evidence from event-related fMRI. *Neuroimage*, 34(3), 1270-1279.



Table 1. Examples of four stimulus sets heard by a single participant

Condition	Lead in	Ambiguity/ Control	Continuation	Sentence- final word
Ambiguous Resolved	The man knew that one more	ACE	might be enough to win the	tennis
Ambiguous Unresolved	The woman hoped that one more	ACE	might be enough to win the	game
Control Resolved	The woman hoped that one more	SPRINT	might be enough to win the	tennis
Control Unresolved	The man knew that one more	SPRINT	might be enough to win the	game
Ambiguous Resolved	Dan looked over all the	ARTICLES	and found that most of them were	broken
Ambiguous Unresolved	Rob went through all the	ARTICLES	and found that most of them were	useless
Control Resolved	Rob went through all the	HAMMERS	and found that most of them were	broken
Control Unresolved	Dan looked over all the	HAMMERS	and found that most of them were	useless
Ambiguous Resolved	The couple thought that this	JAM	was worse than the one on the	motorway
Ambiguous Unresolved	The man heard that this	JAM	was worse than the one on the	television
Control Resolved	The man heard that this	STORM	was worse than the one on the	motorway
Control Unresolved	The couple thought that this	STORM	was worse than the one on the	television
Ambiguous Resolved	His grandfather joked that this	LEEK	was the biggest he had ever	cooked
Ambiguous Unresolved	His uncle claimed that this	LEEK	was the biggest he had ever	found
Control Resolved	His uncle claimed that this	TROUT	was the biggest he had ever	cooked
Control Unresolved	His grandfather joked that this	TROUT	was the biggest he had ever	found



Table 2. Descriptive statistics (Mean (SD)) for key properties of the four key words

Property	Key word		Sentence-final word	
	Ambiguous	Control	Resolved	Unresolved
N	80	80	80	80
Frequency (Log transformed)	1.61 (0.51)	1.3 (0.61)	1.42 (0.64)	1.83 (0.70)
No. syllables	1.28 (0.48)	1.32 (0.47)	1.98 (0.89)	1.98 (0.95)
No. phonemes	3.85 (1.13)	4.05 (1.15)	5.40 (2.08)	5.06 (1.92)

Table 3. Descriptive statistics for cloze proportions across the four sentence conditions shown by key word (ambiguous or control) and sentence-final word response (matching the resolved word or matching the unresolved word).

Key word	Sentence-final word response	Mean Cloze	SD	Median Cloze	Range
Ambiguous (e.g. <i>ACE</i> )	Resolved (e.g. <i>tennis</i> )	0.03	0.06	0.00	0.00-0.34
	Unresolved (e.g. <i>game</i> )	0.09	0.15	0.03	0.00-0.61
Control (e.g. SPRINT)	Resolved	0.06	0.11	0.00	0.00-0.53
	Unresolved	0.08	0.14	0.01	0.00-0.60

46

47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Table 4. Peak voxel locations (in MNI space) and summary statistics from source analysis of the response to ambiguity.

Cluster	Voxels (n)	Region	Coordinates (mm)			Z
			x	y	z	
1	577	Left inferior temporal gyrus	-52	-14	-36	3.41
2	204	Left precentral gyrus	-36	-4	48	3.12
3	281	Right middle temporal gyrus	58	-6	-22	2.75
4	226	Left calcarine sulcus	-12	-100	-8	2.71
		Left calcarine sulcus	-2	-96	4	1.97
5	218	Right middle frontal gyrus	38	46	12	2.56
6	236	Right inferior temporal gyrus	54	-32	-26	2.49
7	76	Right inferior frontal gyrus (pars triangularis)	36	20	26	2.24
8	463	Right inferior frontal gyrus (pars orbitalis)	48	28	-10	2.23
		Right inferior frontal gyrus (pars orbitalis)	36	26	-18	2.19
		Right inferior frontal gyrus (pars triangularis)	54	26	14	2.07
		Right middle frontal gyrus (orbital)	34	56	-2	2.1
9	78	Right superior frontal gyrus (orbital)	26	48	-4	2.01
		Right superior frontal gyrus (orbital)	24	56	-4	1.71
10	69	Left middle frontal gyrus	-36	44	10	2.05
11	27	Left supramarginal gyrus	-58	-44	32	2.02
12	33	Left inferior frontal gyrus (pars opercularis)	-40	16	22	2.01
13	48	Right inferior frontal gyrus (pars opercularis)	42	18	10	1.93
14	33	Right calcarine sulcus	10	-96	2	1.85
15	44	Left inferior temporal gyrus*	-62	-40	-14	1.81
16	47	Right superior temporal gyrus	58	0	-2	1.8
17	34	Left occipital pole	-16	-96	-18	1.75
		Left occipital pole	-32	-92	-18	1.68

Regions are labelled using the AAL atlas (Tzourio-Mazoyer et al., 2002). Activations are thresholded voxel-wise at  $p < .05$  (uncorrected) and cluster-wise at  $k > 25$  voxels. \*Note that this cluster borders the Left middle temporal gyrus and in the Harvard-Oxford atlas is labelled as such.

Table 5. Peak voxel locations (in MNI space) and summary statistics from source analysis of the response to disambiguation.

Cluster	Voxels (n)	Region	Coordinates (mm)			Z
			x	y	z	
1	1357	Right supplementary motor area	12	6	62	2.95
		Right superior frontal gyrus	24	12	60	2.93
		Right precentral gyrus	12	-22	70	2.75
2	387	Left inferior temporal gyrus	-52	-24	-26	2.82
		Left fusiform gyrus	-42	-32	-18	1.74
3	253	Left middle frontal gyrus	-36	22	40	2.35
4	150	Right middle frontal gyrus	34	44	30	2.16
		Right middle frontal gyrus	34	36	28	1.98
5	68	Left lateral occipital cortex	-8	-74	52	2.13
		Left precuneus	-4	-72	44	1.91
6	101	Left inferior frontal gyrus (pars opercularis)	-50	12	22	2.12
7	51	Left precentral gyrus	-36	-16	48	2.09
8	65	Left superior frontal gyrus	-6	36	46	1.97
9	73	Left superior frontal gyrus	-18	10	52	1.97
10	78	Right supramarginal gyrus	44	-32	42	1.96
		Right supramarginal gyrus	50	-30	48	1.96
11	27	Left supramarginal gyrus	-50	-28	28	1.92
12	25	Left supplementary motor area	-4	-4	60	1.9

Regions are labelled using the AAL atlas (Tzourio-Mazoyer et al., 2002). Activations are thresholded voxel-wise at  $p < .05$  (uncorrected) and cluster-wise at  $k > 25$  voxels.

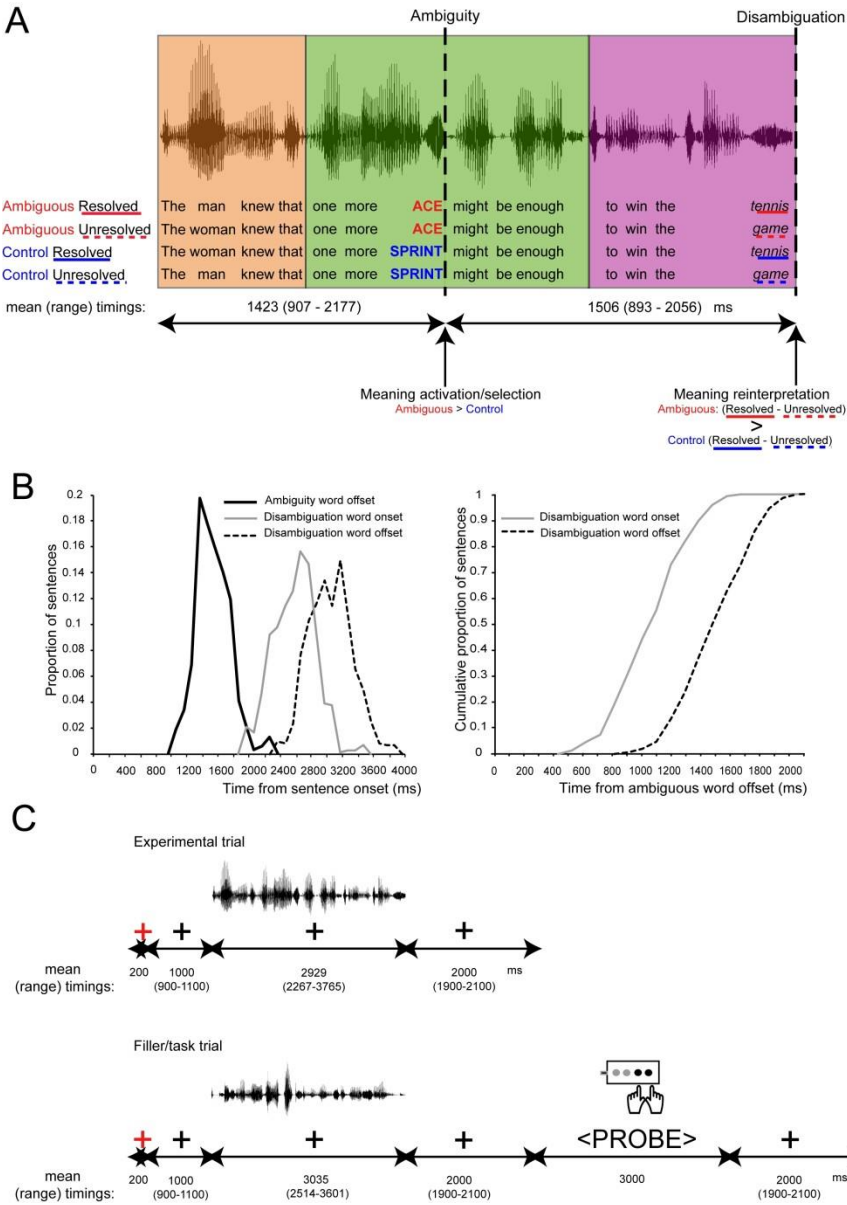


Figure 1. Stimulus and experimental examples and timings. A. Example quartet of the spoken-sentence stimuli showing the four experimental conditions, which were designed to investigate neural processes occurring at two critical time points during semantic ambiguity resolution: (1) Ambiguity and (2) Disambiguation. MEG responses were measured time-locked to the offsets of critical words. At the time of ambiguity, responses to ambiguous words (red) were predicted to be larger than unambiguous control words (blue), reflecting more effortful semantic selection processes. At the time of disambiguation, responses to disambiguating words that resolved the ambiguity to a subordinate meaning (red solid underline) were predicted to be larger than control words that left the ambiguity unresolved (red dotted underline) and control words that completed the unambiguous sentence with each of the words used in the ambiguous sentences (blue solid/dotted underline), reflecting the greater probability of reinterpretation processes. Each sentence was combined from three fragments (highlighted with background colour) from different recordings such that linguistically identical fragments were acoustically identical across conditions, and so that the splice points occurred at least one word before and one word after the ambiguous/control word. B. Frequency distributions of the time durations (ms) between critical words at Ambiguity and at



Disambiguation, shown as proportions across all 640 sentences (i.e. all conditions). Durations are categorised into 100 ms time bins. The left panel displays the distribution of timings of ambiguity word offsets and of disambiguation word onsets and offsets relative to sentence onsets. The right panel shows the cumulative distribution of timings of the onsets and offsets of the disambiguation words relative to the ambiguity word offsets. The offsets of the disambiguation words occur more than 800 ms after ambiguity word offset for all sentences (i.e. at a time beyond the duration of the analysis window for the ambiguity words), and the onsets of the disambiguation words occur more than 800 ms after ambiguity word offsets for 81% of sentences. C. Structure and timings (mean and range) of the components of the experimental trials (top panel) and the filler/task trials (bottom panel).

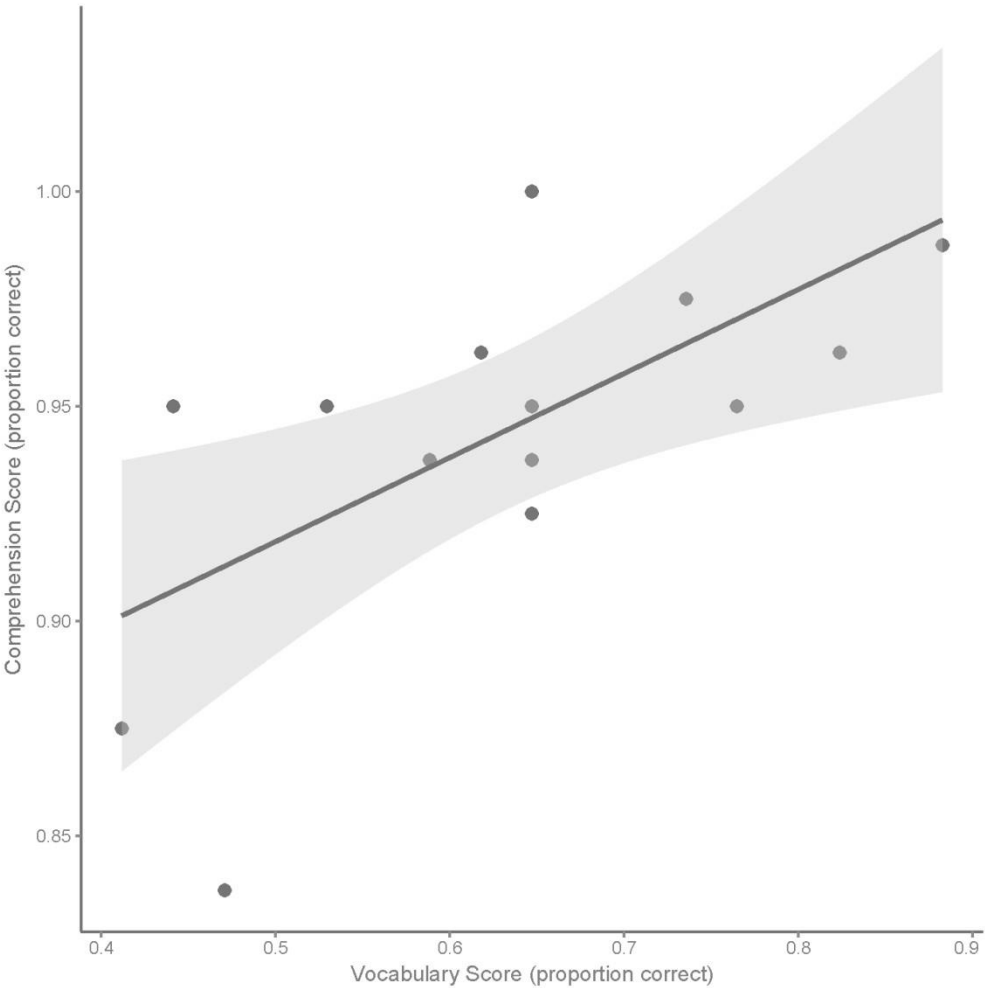


Figure 2. Positive correlation between participants' vocabulary score and their score on the post-MEG comprehension test for the ambiguous resolved sentences. Shaded areas show the 95% Confidence Interval of the regression line.

177x177mm (200 x 200 DPI)

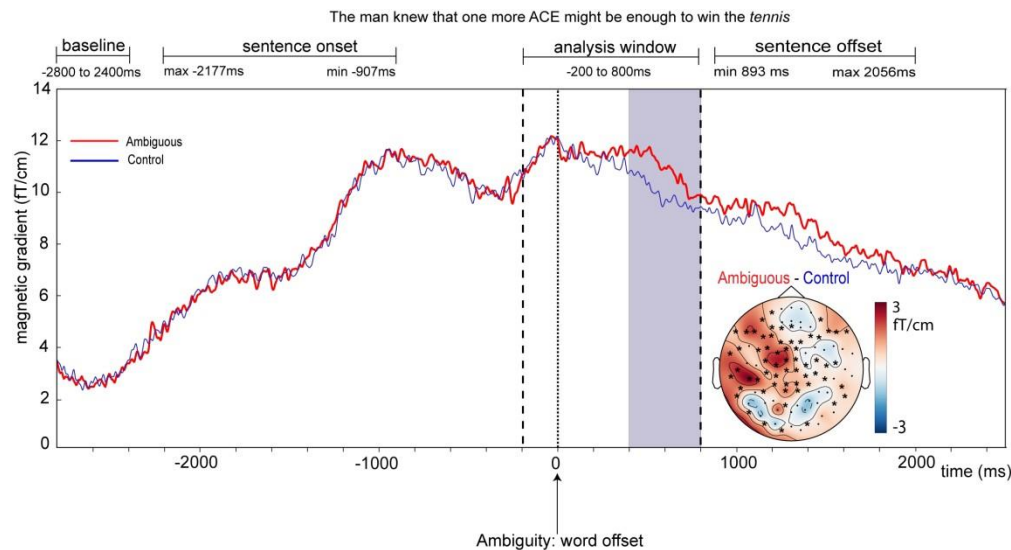


Figure 3. Evoked response at the time of ambiguity for gradiometers. Responses illustrate significantly greater activation for ambiguous (red line) compared to unambiguous control words (blue line) corresponding to a cluster in the data from gradiometer pairs (RMS transformed) beginning approximately 400ms after word offset, which was prominent over left fronto-temporal sensors (analysis time window of -200 to 800 ms relative to word offset). Responses are averaged over all sensors contributing to the significant cluster (highlighted on the topographic plot). Topographic plot shows the distribution over the scalp of the between-condition difference (Ambiguous - Control), averaged over the maximal temporal extent of the cluster (highlighted in purple).

374x203mm (150 x 150 DPI)

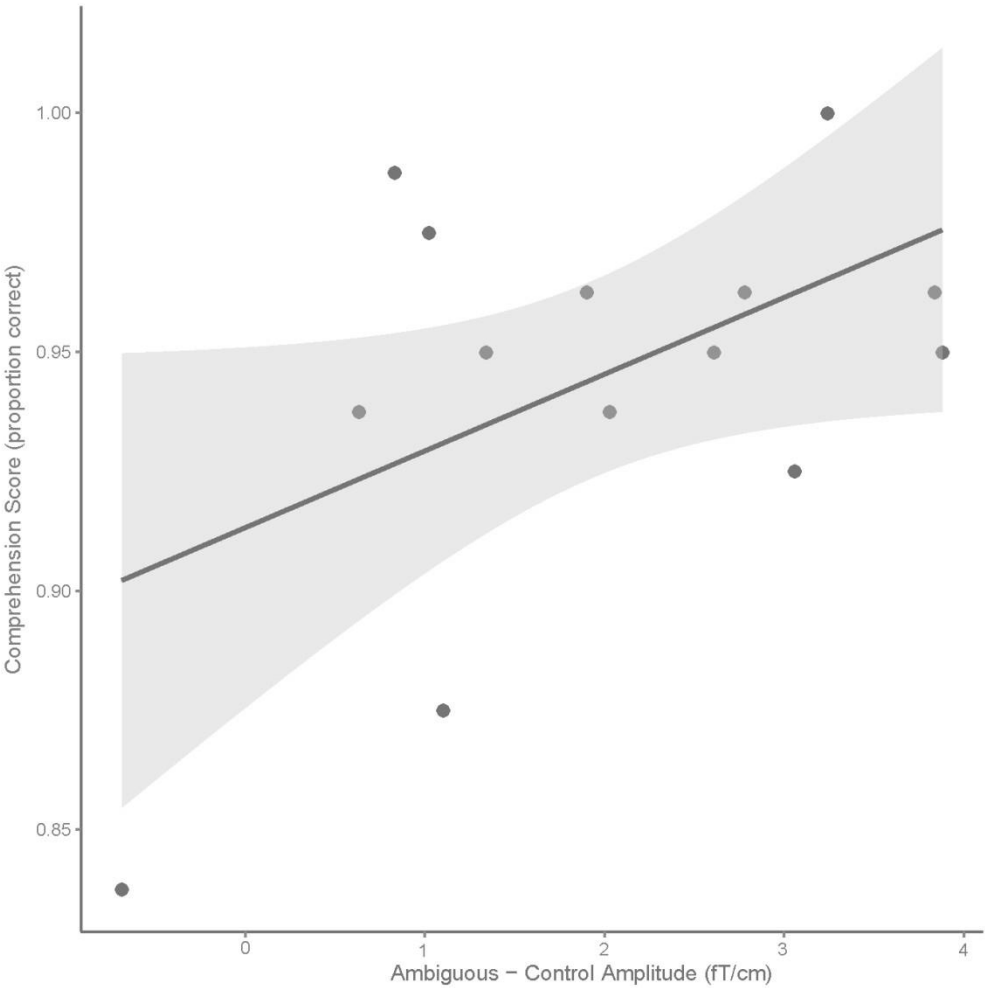


Figure 4. Positive correlation between the amplitude of the MEG effect at the time of ambiguity (Ambiguous - Control) and comprehension scores across participants. Shaded areas show the 95% Confidence Interval of the regression line.

177x177mm (200 x 200 DPI)

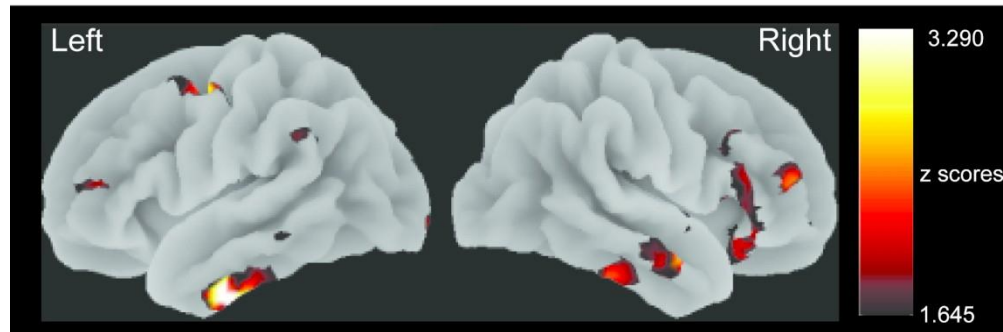


Figure 5. Source localisation of the ambiguity-associated response shown in sensor space analysis in Figure 3. Results show activations displayed at  $p < .05$  (uncorrected) for clarity.

156x52mm (300 x 300 DPI)

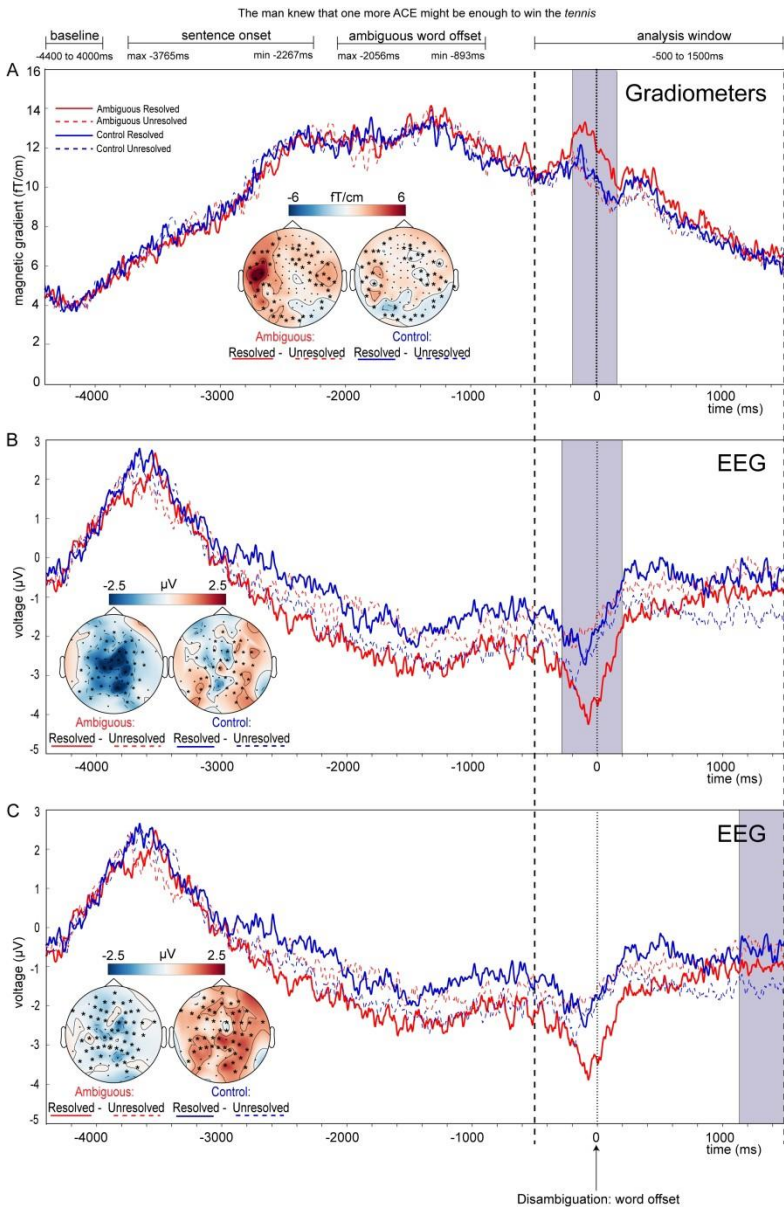


Figure 6. Evoked responses at the time of disambiguation for gradiometers (panel A) and EEG (panels B and C). Responses illustrate marginally significant interactions between Ambiguity and Disambiguation. For the gradiometers (panel A), responses illustrate significantly greater activation for sentence-final words that resolved the ambiguity (red solid line) minus words that left the ambiguity unresolved (red dotted line) compared to the activation difference between identical sentence-final words (blue solid and blue dotted lines) that completed the unambiguous sentences (analysis time window of -500 to 1500 ms relative to word offset). This effect corresponds to a cluster in the data from gradiometer pairs (RMS transformed), which is prominent around word offset and is visually similar to a cluster in the data from EEG (panel B). There is a second cluster for EEG data (panel C), corresponding to a significantly greater difference in activation between the sentence-final words for unambiguous sentences than the difference when these words completed ambiguous sentences. Responses are averaged over all sensors contributing to the significant cluster (highlighted on the topographic plot). Topographic plots show the distribution over the scalp of the between-condition differences (resolved – unresolved), averaged over the maximal temporal extent of the clusters (highlighted in purple), for Ambiguous and Control conditions separately.

375x577mm (150 x 150 DPI)

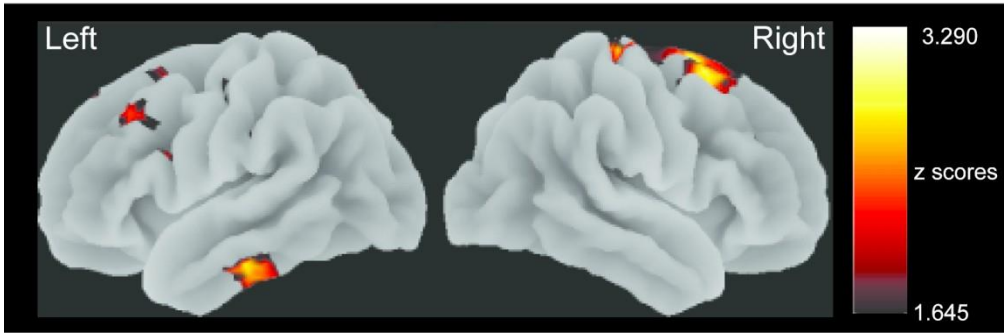


Figure 7. Source localisation of the disambiguation-associated response shown in sensor space analysis in Figure 6. Results show activations displayed at  $p < .05$  (uncorrected) for clarity.  
156x52mm (300 x 300 DPI)